

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:39:11 ; Search time 155.258 Seconds  
(without alignments)  
373.662 Million cell updates/sec

Title: US-10-054-873-7  
Perfect score: 797  
Sequence: 1 MFPTIPLSRLFDNAMLRAHR.....IVEQCCTSICSLYQLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	797	100.0	150	2	AA42861	Aay42861 Chimeric
2	555.5	69.7	107	2	AA42860	Aay42860 hGH-mini-
3	470	59.0	92	2	AA42856	Aay42856 Human gro
4	470	59.0	134	2	AA92265	Aaw92265 Human ant
5	470	59.0	191	5	AB94861	Abg94861 Human gro
6	466	58.5	188	8	AD47330	Adi47330 Plasmid p
7	466	58.5	192	1	AA90129	Aap90129 Human gro
8	466	58.5	192	2	AA92264	Aaw92264 Human ant
9	466	58.5	192	8	AD47320	Adi47320 Plasmid p

10	466	58.5	192	8	ADI47390	Adi47390 Plasmid p
11	466	58.5	192	8	ADI47398	Adi47398 Nmer ampl
12	466	58.5	193	8	ADI47354	Adi47354 Plasmid p
13	466	58.5	206	8	ADI47384	Adi47384 Plasmid p
14	466	58.5	391	8	ADI47363	Adi47363 Plasmid p
15	466	58.5	574	8	ADI47344	Adi47344 Plasmid p
16	466	58.5	576	8	ADI47351	Adi47351 Plasmid p
17	466	58.5	589	8	ADI47365	Adi47365 N+lmer am
18	466	58.5	786	8	ADI47367	Adi47367 Nmer ampl
19	466	58.5	810	8	ADI47388	Adi47388 Amplifica
20	465	58.3	140	1	AAP91041	Aap91041 Human gro
21	465	58.3	191	8	ADR21016	Adr21016 Human mul
22	465	58.3	261	1	AAP91299	Aap91299 Human ner
23	465	58.3	262	2	AAR11740	Aar11740 Human gro
24	465	58.3	310	2	AAR03255	Aar03255 Fusion pr
25	464.5	58.3	191	7	ADF70881	Adf70881 Human gro
26	464	58.2	191	5	ABG31862	Abg31862 Mature hu
27	464	58.2	191	8	ADL89541	Adl89541 Human mod
28	463	58.1	191	5	ABG94860	Abg94860 Human gro
29	463	58.1	191	5	ABG94977	Abg94977 Human gro
30	463	58.1	191	7	ADK41914	Adk41914 Human gro
31	463	58.1	191	8	ADL89532	Adl89532 Human mod
32	463	58.1	191	8	ADL89542	Adl89542 Human mod
33	463	58.1	191	8	ADL89553	Adl89553 Human mod
34	463	58.1	191	8	ADR21180	Adr21180 Human gro
35	463	58.1	213	8	ADQ39291	Adq39291 Human myo
36	463	58.1	794	7	ADF16507	Adf16507 Human alb
37	463	58.1	800	7	ADF16216	Adf16216 Human alb
38	462	58.0	144	2	AAR05313	Aar05313 Segment o
39	462	58.0	191	5	ABG94975	Abg94975 Human gro
40	462	58.0	191	5	ABG94976	Abg94976 Human gro
41	462	58.0	191	7	ADK41912	Adk41912 Human gro
42	462	58.0	191	7	ADK41913	Adk41913 Human gro
43	462	58.0	191	8	ADL89547	Adl89547 Human mod
44	462	58.0	191	8	ADL89554	Adl89554 Human mod
45	462	58.0	191	8	ADL89533	Adl89533 Human mod

#### ALIGNMENTS

##### RESULT 1

AA42861

ID AAY42861 standard; protein; 150 AA.

XX

AC AAY42861;

XX

DT 19-JAN-2000 (first entry)

XX

DE Chimeric protein, SEQ ID 7.

XX

KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;  
KW conformation; chimeric protein; cleavable; recombinant; production;  
KW yield.

XX

OS Synthetic.

OS Homo sapiens.

XX  
 PN WO9950302-A1.  
 XX  
 PD 07-OCT-1999.  
 XX  
 PF 31-MAR-1998; 98WO-CN000052.  
 XX  
 PR 31-MAR-1998; 98WO-CN000052.  
 XX  
 PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.  
 XX  
 PI Gan Z;  
 XX  
 DR WPI; 1999-610839/52.  
 XX  
 PT New chimeric proteins containing human growth hormone fragment, used  
 PT particularly for the production of human insulin.  
 XX  
 PS Claim 14; Page 30-31; 46pp; English.  
 XX  
 CC This sequence represents a chimeric protein, which contains an N-terminal  
 CC fragment of human growth hormone (hGH) of the sequence given in AAY42856,  
 CC a cleavable peptide linker (AAY42857), and a human insulin precursor  
 CC comprising insulin A and B chains (AAY42859). The hGH portion of the  
 CC chimeric protein acts as an intramolecular chaperone (IMC) for the  
 CC insulin precursor, enabling it to fold correctly. The cleavable peptide  
 CC linker has a C-terminal Arg residue which enables the hGH portion of the  
 CC chimeric protein to be removed after folding has taken place. Production  
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can  
 CC provide human insulin with correctly linked cysteine bridges with fewer  
 CC necessary procedural steps, and hence resulting in a higher yield of  
 CC human insulin. The IMC sequences not only protect insulin sequences from  
 CC intracellular degradation by a microorganism host, but also promote the  
 CC folding of the fused insulin precursor, facilitate the solubility of the  
 CC fusion protein and decrease the intermolecular interactions among the  
 CC fusion proteins, thus allowing folding of the fused insulin precursor at  
 CC commercially useful high concentrations. The procedural steps of cyanogen  
 CC bromide cleavage, oxidative sulphytolysis and related purification steps  
 CC can thus be eliminated, along with the use of high concentrations of  
 CC mercaptan or the use of hydrophobic absorbent resins  
 XX  
 SQ Sequence 150 AA;

Query Match 100.0%; Score 797; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-45;  
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDQYEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDQYEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60  
 Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGSHLVEALYLVCGER 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGSHLVEALYLVCGER 120  
 Qy 121 GFFYTPKTRGIVEQCCTSICSLYQLENYCN 150  
 ||||||||||||||||||||||||||||||||||||

## RESULT 2

AAY42860

ID AAY42860 standard; protein; 107 AA.

XX

AC AAY42860;

XX

DT 19-JAN-2000 (first entry)

XX

DE hGH-mini-proinsulin chimeric protein.

XX

KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;  
KW conformation; chimeric protein; cleavable; recombinant; production;  
KW yield.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN WO9950302-A1.

XX

PD 07-OCT-1999.

XX

PF 31-MAR-1998; 98WO-CN000052.

XX

PR 31-MAR-1998; 98WO-CN000052.

XX

PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX

PI Gan Z;

XX

DR WPI; 1999-610839/52.

XX

PT New chimeric proteins containing human growth hormone fragment, used  
PT particularly for the production of human insulin.

XX

PS Claim 13; Page 30; 46pp; English.

XX

CC This sequence represents a chimeric protein, hGH-mini-proinsulin. This  
CC chimeric protein contains an N-terminal fragment of human growth hormone  
CC (hGH) of the sequence given in AAY42855, a cleavable peptide linker  
CC (AAY42857), and a human insulin precursor comprising insulin A and B  
CC chains (AAY42859). The hGH portion of the chimeric protein acts as an  
CC intramolecular chaperone (IMC) for the insulin precursor, enabling it to  
CC fold correctly. The cleavable peptide linker has a C-terminal Arg residue  
CC which enables the hGH portion of the chimeric protein to be removed after  
CC folding has taken place. Production of recombinant human insulin via an  
CC hGH-proinsulin chimeric protein can provide human insulin with correctly  
CC linked cysteine bridges with fewer necessary procedural steps, and hence  
CC resulting in a higher yield of human insulin. The IMC sequences not only  
CC protect insulin sequences from intracellular degradation by a  
CC microorganism host, but also promote the folding of the fused insulin  
CC precursor, facilitate the solubility of the fusion protein and decrease  
CC the intermolecular interactions among the fusion proteins, thus allowing  
CC folding of the fused insulin precursor at commercially useful high  
CC concentrations. The procedural steps of cyanogen bromide cleavage,



CC oxidative sulphitolysis and related purification steps can thus be  
CC eliminated, along with the use of high concentrations of mercaptan or the  
CC use of hydrophobic absorbent resins

XX

SQ Sequence 107 AA;

Query Match 69.7%; Score 555.5; DB 2; Length 107;  
Best Local Similarity 71.3%; Pred. No. 1.1e-29;  
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60  
|||||

Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNP----- 49

Qy 61 TPSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLGSHLVEALYLVCGER 120  
|||||

Db 50 -----LGTGPRFVNQHLGSHLVEALYLVCGER 77

Qy 121 GFFYTPKTRGIVEQCCTSICSISLYQLENYCN 150  
|||||

Db 78 GFFYTPKTRGIVEQCCTSICSISLYQLENYCN 107

### RESULT 3

AA42856

ID AA42856 standard; protein; 92 AA.

XX

AC AA42856;

XX

DT 19-JAN-2000 (first entry)

XX

DE Human growth hormone (hGH) N-terminal fragment #2.

XX

KW Growth hormone; chaperone; intramolecular; insulin; precursor; folding;

KW conformation; chimeric protein; cleavable; recombinant; production;

KW yield.

XX

OS Homo sapiens.

XX

PN WO9950302-A1.

XX

PD 07-OCT-1999.

XX

PF 31-MAR-1998; 98WO-CN000052.

XX

PR 31-MAR-1998; 98WO-CN000052.

XX

PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX

PI Gan Z;

XX

DR WPI; 1999-610839/52.

XX

PT New chimeric proteins containing human growth hormone fragment, used

PT particularly for the production of human insulin.

XX

PS Claim 5; Page 28; 46pp; English.

XX  
 CC This sequence represents an N-terminal fragment of human growth hormone  
 CC (hGH) which is a component of a chimeric protein (AAY42861) which also  
 CC contains a human insulin precursor (AAY42859). The hGH portion of the  
 CC chimeric protein acts as an intramolecular chaperone (IMC) for the  
 CC insulin precursor, enabling it to fold correctly. A cleavable peptide  
 CC linker with a C-terminal Arg residue (AAY42857) enables the hGH portion  
 CC of the chimeric protein to be removed after folding has taken place.  
 CC Production of recombinant human insulin via an hGH-proinsulin chimeric  
 CC protein can provide human insulin with correctly linked cysteine bridges  
 CC with fewer necessary procedural steps, and hence resulting in a higher  
 CC yield of human insulin. The IMC sequences not only protect insulin  
 CC sequences from intracellular degradation by a microorganism host, but  
 CC also promote the folding of the fused insulin precursor, facilitate the  
 CC solubility of the fusion protein and decrease the intermolecular  
 CC interactions among the fusion proteins, thus allowing folding of the  
 CC fused insulin precursor at commercially useful high concentrations. The  
 CC procedural steps of cyanogen bromide cleavage, oxidative sulphytolysis  
 CC and related purification steps can thus be eliminated, along with the use  
 CC of high concentrations of mercaptan or the use of hydrophobic absorbent  
 CC resins  
 XX  
 SQ Sequence 92 AA;

Query Match 59.0%; Score 470; DB 2; Length 92;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-24;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
 Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92  
 ||||||||||||||||||||||||||||  
 Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 4  
 AAW92265

ID AAW92265 standard; protein; 134 AA.  
 XX  
 AC AAW92265;  
 XX  
 DT 08-JUN-1999 (first entry)  
 XX  
 DE Human anti-angiogenic peptide 16K hGH Met-1Pro133.  
 XX  
 KW Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;  
 KW growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;  
 KW placental vascularisation; pregnancy; treatment; angiogenic disease;  
 KW tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;  
 KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;  
 KW wound healing; proliferative retinopathy; macular degeneration; trachoma;  
 KW granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;  
 KW psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;  
 KW ulcer; leukaemia; reproductive disorder; contraceptive agent;  
 KW gene therapy; pre-eclampsia; intrauterine growth retardation;

KW placental dysfunction.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9851323-A1.  
 XX  
 PD 19-NOV-1998.  
 XX  
 PF 12-MAY-1998; 98WO-US009691.  
 XX  
 PR 13-MAY-1997; 97US-0046394P.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Weiner RI, Martial JA, Struman I, Taylor R;  
 XX  
 DR WPI; 1999-045192/04.  
 DR N-PSDB; AAX01707.  
 XX  
 PT New anti-angiogenic peptides - comprise N-terminal fragments of human  
 PT placental lactogen, human growth hormone, growth hormone variant or human  
 PT prolactin.  
 XX  
 PS Claim 4; Page 49-50; 87pp; English.  
 XX  
 CC This invention describes novel human anti-angiogenic peptides derived  
 CC from 10 to 150 consecutive amino acids selected from the N-terminal end  
 CC of human placental lactogen (hPL), human growth hormone (hGH), growth  
 CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit  
 CC capillary endothelial cell proliferation and organisation (ii) inhibit  
 CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at  
 CC least one specific receptor which does not bind an intact full length  
 CC hGH, hPL, prolactin or hGH-V. The invention also describes a method for  
 CC diagnosing a probable abnormality of placental vascularisation during  
 CC pregnancy. The peptides can be used for treating an angiogenic disease in  
 CC a subject, for inhibiting tumour formation or growth in a patient or for  
 CC modulating vascularisation of a patient's placenta. In particular, the  
 CC peptides can be used for preventing or treating e.g. malignant tumours,  
 CC angiofibroma, arteriovenous malformation, arthritic such as rheumatoid  
 CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,  
 CC delayed wound healing, proliferative retinopathy such as diabetic  
 CC retinopathy, macular degeneration, granulations such as those occurring  
 CC in haemophilic joints, inappropriate vascularisation in wound healing  
 CC such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular  
 CC tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis,  
 CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,  
 CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,  
 CC leukaemia, and reproductive disorders such as follicular and luteal cysts  
 CC and choriocarcinoma. They can also be used as contraceptive agents. DNA  
 CC encoding the peptides can be used in gene therapy. The measurement of  
 CC abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL  
 CC can be used in assays for impairment of vascular development associated  
 CC with pre-eclampsia, intrauterine growth retardation, and placental  
 CC dysfunction  
 XX  
 SQ Sequence 134 AA;

Query Match 59.0%; Score 470; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 6.1e-24;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
          ||||||||||||||||||||||||
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
```

RESULT 5

ABG94861

ID ABG94861 standard; protein; 191 AA.

XX

AC ABG94861;

XX

DT 03-DEC-2002 (first entry)

XX

DE Human growth hormone mutant hPRL (111-129).

XX

KW Growth hormone; placental lactogen; prolactin; active domain; hGH;

KW structure-function relationship; segment-substituted polypeptide; mutant;

KW mutein.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN US6428954-B1.

XX

PD 06-AUG-2002.

XX

PF 06-JUN-1995; 95US-00483039.

XX

PR 28-OCT-1988; 88US-00264611.

PR 26-OCT-1989; 89US-00428066.

PR 27-APR-1992; 92US-00875204.

PR 13-OCT-1992; 92US-00960227.

PR 02-FEB-1994; 94US-00190723.

XX

PA (GETH ) GENENTECH INC.

XX

PI Wells JA, Cunningham BC;

XX

DR WPI; 2002-696875/75.

XX

PT Identifying active domains within cloned polypeptides of known amino acid

PT sequence by substituting analog segments into the parent polypeptide is

PT useful to determine the relationship between structure and function.

XX

PS Example 1; Page; 86pp; English.

XX

CC The invention relates to identifying an unknown active domain in a region

CC of known amino acid sequence in a parent polypeptide e.g. human growth

CC hormone (hGH) which has been cloned and has a pre-identified biological

CC activity, where the active domain interacts with a target when the parent  
 CC polypeptide is in its native-folded form and the interaction is  
 CC responsible for the biological activity comprising: (a) comparing the  
 CC amino acid sequence or polypeptide structure in the region of known amino  
 CC acid sequence of hGH with the amino acid sequence or polypeptide  
 CC structure in a region of known amino acid sequence of an analogue  
 CC polypeptide (e.g. prolactin, placental lactogen or porcine growth  
 CC hormone) which has at least 15% homology with hGH alpha-carbon  
 CC coordinates within about 2-3.5 angstroms of hGH alpha-carbon coordinates  
 CC for about 60% of the analogue sequence, where any interaction of the  
 CC analogue with the target is different from target interaction with hGH;  
 CC (b) substituting DNA encoding an analogous polypeptide segment from the  
 CC analogue into DNA encoding the full length hGH, and expressing a segment-  
 CC substituted polypeptide; (c) contacting the segment-substituted  
 CC polypeptide with the target to determine interaction; (d) repeating steps  
 CC (b) and (c) with a second analogous polypeptide segment; and (e)  
 CC comparing the difference between activity of the first and second segment  
 CC -substituted polypeptides as an indication of the location of the unknown  
 CC active domain in hGH. The method is useful for determining the  
 CC relationship between structure and function of known polypeptide  
 CC sequences. The present sequence is that of human growth hormone mutant  
 CC substituted with residues from an hGH analogue (prolactin, placental  
 CC lactogen or porcine growth hormone). Note: The present sequence is not  
 CC shown in the specification but was created by the indexer using the  
 CC mature hGH sequence and information contained in the specification  
 XX  
 SQ Sequence 191 AA;

Query Match 59.0%; Score 470; DB 5; Length 191;  
 Best Local Similarity 69.2%; Pred. No. 8.2e-24;  
 Matches 101; Conservative 8; Mismatches 19; Indels 18; Gaps 3;

```
Qy      2 FFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          |||
Db      1 FFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 60

Qy      62 PSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----H 108
          |||
Db      61 PSNREETQQKSNLELLRLISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVVDILEQLKR 119

Qy      109 LVEALYLVCGERGFFYTPKTRGIVEQ 134
          |:| | |: : : |:| | :|
Db      120 LIEGLMLILSDG----SPRTGQIFKQ 141
```

RESULT 6

ADI47330

ID ADI47330 standard; protein; 188 AA.

XX

AC ADI47330;

XX

DT 22-APR-2004 (first entry)

XX

DE Plasmid p0A11A1 amino acid sequence SEQ ID NO:18.

XX

KW multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

KW therapeutic protein.

XX

OS Synthetic.

XX

PN WO2004007687-A2.

XX

PD 22-JAN-2004.

XX

PF 16-JUL-2003; 2003WO-US022216.

XX

PR 16-JUL-2002; 2002US-0396466P.

XX

PA (BUSS/) BUSSELL S.

XX

PI Bussell S;

XX

DR WPI; 2004-122926/12.

DR N-PSDB; ADI47329.

XX

PT Multimer assembly of DNA sequences comprising an amplification cassette  
PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
PT terminus and 3' RPM at its 3' terminus.

XX

PS Example 2; SEQ ID NO 18; 163pp; English.

XX

CC The present invention describes a multimer assembly of DNA sequences (I)  
CC comprising at least one amplification cassette (AC) having at least one  
CC monomer sequence whose polymerisation is desired, and a 5' restriction  
CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
CC factor, a DNA replication factor, an activator, a chaperonin, or an  
CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
CC granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory  
CC factor, tumour necrosis factor, lymphotoxin, platelet-derived growth  
CC factor, fibroblast growth factors, vascular endothelial cell growth  
CC factor, epidermal growth factor, transforming growth factor-beta,  
CC transforming growth factor-alpha, thrombopoietin, stem cell factor, ...  
CC oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth  
CC factor, macrophage migration inhibiting factor, endostatin, or  
CC angiostatin. The present sequence is used in the exemplification of the  
CC present invention.

XX

SQ Sequence 188 AA;

Query Match 58.5%; Score 466; DB 8; Length 188;

Best Local Similarity 70.5%; Pred. No. 1.5e-23;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
 |||||  
 Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60  
 Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110  
 ||||| : | : |  
 Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119  
 Qy 111 EALYLVCG--ERGFFYTPKTRGIVEQ 134  
 | : : | | : | : |  
 Db 120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

# RESULT 7

AAP90129

ID AAP90129 standard; protein; 192 AA.

XX

AC AAP90129;

XX

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 06-FEB-1996 (revised)

DT 01-NOV-1989 (first entry)

XX

DE Human growth hormone.

XX

KW Human growth hormone; fusion protein; recombinant vector.

XX

OS Homo sapiens; (Human).

XX

PN JP01144981-A.

XX

PD 07-JUN-1989.

XX

PF 02-DEC-1987; 87JP-00304937.

XX

PR 02-DEC-1987; 87JP-00304937.

XX

PA (WAKT ) WAKUNAGA SEIYAKU KK.

XX

DR WPI; 1989-209284/29.

DR N-PSDB; AAN90269.

XX

PT Recombinant vector contg. fused protein aminoacid coding - composed of  
 PT growth hormone or its polypeptide deriv. and foreign protein.

XX

PS Disclosure; Fig 1; 19pp; Japanese.

XX

CC The invention consists of a vector contg. a fusion protein which is  
 CC formed by ligating, downstream of a promoter, hGH or a deriv. (pref.  
 CC formed by substn. of Met-14 with Leu) and a foreign protein. Stability  
 CC of the vector in the host is greatly increased so the protein yield is  
 CC higher. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-  
 CC 2003 to standardise OS field)

XX

SQ Sequence 192 AA;

Query Match 58.5%; Score 466; DB 1; Length 192;  
 Best Local Similarity 70.5%; Pred. No. 1.5e-23;  
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60

Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110  
 ||||||||||||||||||||||||||||| | | : | : |  
 Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy 111 EALYLVCG--ERGFFYTPKTRGIVEQ 134  
 | : : | | | : | : | : |  
 Db 120 EGIQTLMGRLDGD---SPRTGQIFKQ 142

RESULT 8

AAW92264

ID AAW92264 standard; protein; 192 AA.

XX

AC AAW92264;

XX

DT 08-JUN-1999 (first entry)

XX

DE Human anti-angiogenic peptide hGH Met-1Phe191.

XX

KW Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;  
 KW growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;  
 KW placental vascularisation; pregnancy; treatment; angiogenic disease;  
 KW tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;  
 KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;  
 KW wound healing; proliferative retinopathy; macular degeneration; trachoma;  
 KW granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;  
 KW psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;  
 KW ulcer; leukaemia; reproductive disorder; contraceptive agent;  
 KW gene therapy; pre-eclampsia; intrauterine growth retardation;  
 KW placental dysfunction.

XX

OS Homo sapiens.

XX

PN WO9851323-A1.

XX

PD 19-NOV-1998.

XX

PF 12-MAY-1998; 98WO-US009691.

XX

PR 13-MAY-1997; 97US-0046394P.

XX

PA (REGC ) UNIV CALIFORNIA.

XX

PI Weiner RI, Martial JA, Struman I, Taylor R;

XX

DR WPI; 1999-045192/04.

DR N-PSDB; AAX01706.

XX

PT New anti-angiogenic peptides - comprise N-terminal fragments of human





XX  
 AC ADI47320;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE Plasmid p0A0 amino acid sequence SEQ ID NO:8.  
 XX  
 KW multimer assembly; DNA sequence; amplification cassette;  
 KW monomer sequence; restriction pair member; diagnostic protein;  
 KW therapeutic protein.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2004007687-A2.  
 XX  
 PD 22-JAN-2004.  
 XX  
 PF 16-JUL-2003; 2003WO-US022216.  
 XX  
 PR 16-JUL-2002; 2002US-0396466P.  
 XX  
 PA (BUSS/) BUSSELL S.  
 XX  
 PI Bussell S;  
 XX  
 DR WPI; 2004-122926/12.  
 DR N-PSDB; ADI47319.  
 XX  
 PT Multimer assembly of DNA sequences comprising an amplification cassette  
 PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
 PT terminus and 3' RPM at its 3' terminus.  
 XX  
 PS Example 1; SEQ ID NO 8; 163pp; English.  
 XX  
 CC The present invention describes a multimer assembly of DNA sequences (I)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerisation is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
 CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
 CC factor, a DNA replication factor, an activator, a chaperonin, or an  
 CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
 CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
 CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
 CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
 CC granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory  
 CC factor, tumour necrosis factor, lymphotoxin, platelet-derived growth  
 CC factor, fibroblast growth factors, vascular endothelial cell growth  
 CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or

CC angiostatin. The present sequence is used in the exemplification of the  
CC present invention.  
XX  
SQ Sequence 192 AA;

Query Match 58.5%; Score 466; DB 8; Length 192;  
Best Local Similarity 70.5%; Pred. No. 1.5e-23;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          |||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCS-----HLV 110
          |||
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy    111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | : | | : |
Db    120 EGIQTLMGRLDG---SPRTGQIFKQ 142
```

RESULT 10

ADI47390

ID ADI47390 standard; protein; 192 AA.

XX

AC ADI47390;

XX

DT 22-APR-2004 (first entry)

XX

DE Plasmid p0A51A amino acid sequence SEQ ID NO:78.

XX

KW multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

KW therapeutic protein.

XX

OS Synthetic.

XX

PN WO2004007687-A2.

XX

PD 22-JAN-2004.

XX

PF 16-JUL-2003; 2003WO-US022216.

XX

PR 16-JUL-2002; 2002US-0396466P.

XX

PA (BUSS/) BUSSELL S.

XX

PI Bussell S;

XX

DR WPI; 2004-122926/12.

DR P-PSDB; ADI47389.

XX

PT Multimer assembly of DNA sequences comprising an amplification cassette

PT having monomer sequences and 5' restriction pair member (RPM) at its 5'

PT terminus and 3' RPM at its 3' terminus.

XX

PS Example 12; SEQ ID NO 78; 163pp; English.

XX

CC The present invention describes a multimer assembly of DNA sequences (I) comprising at least one amplification cassette (AC) having at least one monomer sequence whose polymerisation is desired, and a 5' restriction pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and one or more of following: (a) 3'-terminal cassette comprising 3' specific sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal cassette comprising 5' specific sequence and 3' RPM site fused to a 5' RPM site of AC. (I) can be used for expressing a diagnostic protein or therapeutic protein. In (I), the diagnostic protein and therapeutic protein is a cytokine, a growth factor, a hormone, a receptor, a receptor ligand, an enzyme, an inhibitor, a transcription factor, a translation factor, a DNA replication factor, an activator, a chaperonin, or an antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta, IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin, colony-stimulating factor-1, granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory factor, tumour necrosis factor, lymphotoxin, platelet-derived growth factor, fibroblast growth factors, vascular endothelial cell growth factor, epidermal growth factor, transforming growth factor-beta, transforming growth factor-alpha, thrombopoietin, stem cell factor, oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth factor, macrophage migration inhibiting factor, endostatin, or angiostatin. The present sequence is used in the exemplification of the present invention.

XX

SQ Sequence 192 AA;

Query Match 58.5%; Score 466; DB 8; Length 192;  
Best Local Similarity 70.5%; Pred. No. 1.5e-23;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
QY      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60
          |||
Db       1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

QY      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPREFVNQHLCGS-----HLV 110
          |||
Db       61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

QY      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | :| | :|
Db       120 EGIQTLMGRLEDG---SPRTGQIFKQ 142
```

RESULT 11

ADI47398

ID ADI47398 standard; protein; 192 AA.

XX

AC ADI47398;

XX

DT 22-APR-2004 (first entry)

XX

DE Nmer amplification cassette amino acid sequence SEQ ID NO:86.

XX

KW multimer assembly; DNA sequence; amplification cassette;  
 KW monomer sequence; restriction pair member; diagnostic protein;  
 KW therapeutic protein.  
 XX  
 OS Synthetic.  
 XX  
 PN WO2004007687-A2.  
 XX  
 PD 22-JAN-2004.  
 XX  
 PF 16-JUL-2003; 2003WO-US022216.  
 XX  
 PR 16-JUL-2002; 2002US-0396466P.  
 XX  
 PA (BUSS/) BUSSELL S.  
 XX  
 PI Bussell S;  
 XX  
 DR WPI; 2004-122926/12.  
 DR P-PSDB; ADI47397.  
 XX  
 PT Multimer assembly of DNA sequences comprising an amplification cassette  
 PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
 PT terminus and 3' RPM at its 3' terminus.  
 XX  
 PS Claim 115; SEQ ID NO 86; 163pp; English.  
 XX  
 CC The present invention describes a multimer assembly of DNA sequences (I)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerisation is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
 CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
 CC factor, a DNA replication factor, an activator, a chaperonin, or an  
 CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
 CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
 CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
 CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
 CC granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory  
 CC factor, tumour necrosis factor, lymphotoxin, platelet-derived growth  
 CC factor, fibroblast growth factors, vascular endothelial cell growth  
 CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or  
 CC angiostatin. The present sequence is used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 192 AA;

Query Match 58.5%; Score 466; DB 8; Length 192;  
 Best Local Similarity 70.5%; Pred. No. 1.5e-23;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

QY      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

QY      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
      ||||||||||||||||||||||||||||| | | : | : |
Db      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

QY      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
      | : : | | | :|:| | :|
Db      120 EGIQTLMGRLLEDG---SPRTGQIFKQ 142

```

# RESULT 12

ADI47354

ID ADI47354 standard; protein; 193 AA.

XX

AC ADI47354;

XX

DT 22-APR-2004 (first entry)

XX

DE Plasmid p0A31A amino acid sequence SEQ ID NO:42.

XX

KW multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

KW therapeutic protein.

XX

OS Synthetic.

XX

PN WO2004007687-A2.

XX

PD 22-JAN-2004.

XX

PF 16-JUL-2003; 2003WO-US022216.

XX

PR 16-JUL-2002; 2002US-0396466P.

XX

PA (BUSS/) BUSSELL S.

XX

PI Bussell S;

XX

DR WPI; 2004-122926/12.

DR N-PSDB; ADI47353.

XX

PT Multimer assembly of DNA sequences comprising an amplification cassette having monomer sequences and 5' restriction pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus.

XX

PS Example 7; SEQ ID NO 42; 163pp; English.

XX

CC The present invention describes a multimer assembly of DNA sequences (I) comprising at least one amplification cassette (AC) having at least one monomer sequence whose polymerisation is desired, and a 5' restriction pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and one or more of following: (a) 3'-terminal cassette comprising 3' specific







				:	:	
Db	61	TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE	119			
Qy	111	EALYLVCG--ERGFFYTPKTRGIVEQ	134			
		: :       : :    :				
Db	120	EGIQTLMGRLEDG---SPRTGQIFKQ	142			

# RESULT 14

ADI47363

ID ADI47363 standard; protein; 391 AA.

XX

AC ADI47363;

XX

DT 22-APR-2004 (first entry)

XX

DE Plasmid p0A31F2 insert amino acid sequence SEQ ID NO:51.

XX

KW multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

KW therapeutic protein.

XX

OS Synthetic.

XX

PN WO2004007687-A2.

XX

PD 22-JAN-2004.

XX

PF 16-JUL-2003; 2003WO-US022216.

XX

PR 16-JUL-2002; 2002US-0396466P.

XX

PA (BUSS/) BUSSELL S.

XX

PI Bussell S;

XX

DR WPI; 2004-122926/12.

DR N-PSDB; ADI47362.

XX

PT Multimer assembly of DNA sequences comprising an amplification cassette

PT having monomer sequences and 5' restriction pair member (RPM) at its 5'

PT terminus and 3' RPM at its 3' terminus.

XX

PS Example 8; SEQ ID NO 51; 163pp; English.

XX

CC The present invention describes a multimer assembly of DNA sequences (I)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerisation is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
 CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
 CC factor, a DNA replication factor, an activator, a chaperonin, or an



PA (BUSS/) BUSSELL S.

XX

PI Bussell S;

XX

DR WPI; 2004-122926/12.

DR N-PSDB; ADI47343.

XX

PT Multimer assembly of DNA sequences comprising an amplification cassette  
PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
PT terminus and 3' RPM at its 3' terminus.

XX

PS Claim 67; SEQ ID NO 32; 163pp; English.

XX

CC The present invention describes a multimer assembly of DNA sequences (I)  
CC comprising at least one amplification cassette (AC) having at least one  
CC monomer sequence whose polymerisation is desired, and a 5' restriction  
CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
CC factor, a DNA replication factor, an activator, a chaperonin, or an  
CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
CC granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory  
CC factor, tumour necrosis factor, lymphotoxin, platelet-derived growth  
CC factor, fibroblast growth factors, vascular endothelial cell growth  
CC factor, epidermal growth factor, transforming growth factor-beta,  
CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
CC oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth  
CC factor, macrophage migration inhibiting factor, endostatin, or  
CC angiostatin. The present sequence is used in the exemplification of the  
CC present invention.

XX

SQ Sequence 574 AA;

Query Match 58.5%; Score 466; DB 8; Length 574;

Best Local Similarity 70.5%; Pred. No. 3.7e-23;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          |||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
          |||
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy    111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | :|:| | :|
Db    120 EGIQTLMGRLEDG---SPRTGQIFKQ 142
```

Search completed: February 11, 2005, 18:14:54  
Job time : 156.258 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:04:56 ; Search time 39.5756 Seconds  
 (without alignments)  
 282.936 Million cell updates/sec

Title: US-10-054-873-7  
 Perfect score: 797  
 Sequence: 1 MFPTIPLSRLFDNAMLRAHR.....IVEQCCTSICSLYQLENYCN 150

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	466	58.5	192	1	US-08-093-383-1	Sequence 1, Appli
2	463	58.1	242	4	US-09-949-016-8660	Sequence 8660, Ap
3	463	58.1	242	4	US-09-949-016-8661	Sequence 8661, Ap
4	463	58.1	242	4	US-09-949-016-8662	Sequence 8662, Ap
5	463	58.1	242	4	US-09-949-016-8663	Sequence 8663, Ap
6	463	58.1	242	4	US-09-949-016-8664	Sequence 8664, Ap
7	461	57.8	191	3	US-09-284-878-5	Sequence 5, Appli
8	461	57.8	191	4	US-09-462-941-1	Sequence 1, Appli
9	461	57.8	194	2	US-08-383-621-4	Sequence 4, Appli
10	461	57.8	194	3	US-08-459-906-4	Sequence 4, Appli
11	461	57.8	217	3	US-08-589-028-10	Sequence 10, Appli

12	461	57.8	217	3	US-08-784-582-10	Sequence 10, Appl
13	461	57.8	217	3	US-08-785-271-10	Sequence 10, Appl
14	461	57.8	217	3	US-08-759-628-11	Sequence 11, Appl
15	461	57.8	217	3	US-09-284-878-1	Sequence 1, Appli
16	461	57.8	217	4	US-09-929-918-9	Sequence 9, Appli
17	461	57.8	241	3	US-09-424-620B-25	Sequence 25, Appl
18	461	57.8	245	4	US-09-280-030-66	Sequence 66, Appl
19	461	57.8	274	3	US-08-784-582-71	Sequence 71, Appl
20	461	57.8	360	3	US-08-784-582-73	Sequence 73, Appl
21	461	57.8	448	4	US-09-916-229A-2	Sequence 2, Appli
22	460	57.7	191	4	US-09-554-451-1	Sequence 1, Appli
23	460	57.7	198	4	US-09-949-016-8650	Sequence 8650, Ap
24	460	57.7	198	4	US-09-949-016-8651	Sequence 8651, Ap
25	460	57.7	198	4	US-09-949-016-8652	Sequence 8652, Ap
26	460	57.7	198	4	US-09-949-016-8653	Sequence 8653, Ap
27	460	57.7	198	4	US-09-949-016-8654	Sequence 8654, Ap
28	455	57.1	191	3	US-09-465-461-1	Sequence 1, Appli
29	455	57.1	191	4	US-09-554-451-3	Sequence 3, Appli
30	455	57.1	217	1	US-08-187-756C-4	Sequence 4, Appli
31	455	57.1	217	1	US-08-469-486-51	Sequence 51, Appl
32	455	57.1	217	2	US-08-469-658-51	Sequence 51, Appl
33	455	57.1	217	2	US-08-710-324A-4	Sequence 4, Appli
34	455	57.1	217	4	US-09-411-657-4	Sequence 4, Appli
35	454	57.0	400	4	US-09-420-819-37	Sequence 37, Appl
36	454	57.0	401	4	US-09-420-819-36	Sequence 36, Appl
37	448	56.2	191	3	US-08-800-215C-18	Sequence 18, Appl
38	446	56.0	191	3	US-08-800-215C-16	Sequence 16, Appl
39	446	56.0	191	3	US-08-800-215C-20	Sequence 20, Appl
40	365.5	45.9	176	3	US-08-791-728-1	Sequence 1, Appli
41	365.5	45.9	176	3	US-08-990-774-1	Sequence 1, Appli
42	359.5	45.1	176	3	US-08-791-728-2	Sequence 2, Appli
43	359.5	45.1	176	3	US-08-990-774-2	Sequence 2, Appli
44	343	43.0	168	6	5424199-3	Patent No. 5424199
45	343	43.0	168	6	5424199-3	Patent No. 5424199

#### ALIGNMENTS

#### RESULT 1

US-08-093-383-1

; Sequence 1, Application US/08093383

; Patent No. 5489529

; GENERAL INFORMATION:

; APPLICANT: DeBoer, Herman A.

; APPLICANT: Heyneker, Herbert L.

; APPLICANT: Seeburg, Peter H.

; TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,383
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/619827
; FILING DATE: 28-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/198824
; FILING DATE: 05-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/632361
; FILING DATE: 19-JUL-1984
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/303687
; FILING DATE: 18-SEP-1981
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnston, Sean A.
; REGISTRATION NUMBER: P35,910
; REFERENCE/DOCKET NUMBER: 46C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3562
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-093-383-1

```

```

Query Match          58.5%; Score 466; DB 1; Length 192;
Best Local Similarity 70.5%; Pred. No. 1.9e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
        ||||||||||||||||||||||||||||||||| | | : | : |
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy    111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
        | : : | | | : | : | | : |
Db    120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

RESULT 2

```

US-09-949-016-8660
; Sequence 8660, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8660
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8660

```

```

Query Match          58.1%; Score 463; DB 4; Length 242;
Best Local Similarity 65.2%; Pred. No. 5.5e-42;
Matches 103; Conservative 8; Mismatches 27; Indels 20; Gaps 4;

```

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 115

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||||||  | | : | :  | |
Db      116  PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 174

Qy      112  ALYLVCGERGFFYTPK-----TRGIVEQCCTSICSLY 143
          : : | |  |  | : : | |  | | :
Db      175  GIQTLMGVRVAPGVNPGAPLTLRAVLEKHC---CPLF 209

```

# RESULT 3

```

US-09-949-016-8661
; Sequence 8661, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```



; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8661  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8661

Query Match 58.1%; Score 463; DB 4; Length 242;  
Best Local Similarity 65.2%; Pred. No. 5.5e-42;  
Matches 103; Conservative 8; Mismatches 27; Indels 20; Gaps 4;

```
Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      56  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 115

Qy      62  PSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||||||  | | : | :  | |
Db     116  PSNREETQQKSNLELLRLISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 174

Qy     112  ALYLVCGERGFFYTPK-----TRGIVEQCCTSICSLY 143
          : : | |  |  | : : | |  | | :
Db     175  GIQTLMGVRVAPGVPNPGAPLTLRAVLEKHC---CPLF 209
```

#### RESULT 4

US-09-949-016-8662  
; Sequence 8662, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES  
THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8662  
; LENGTH: 242  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-8662

Query Match 58.1%; Score 463; DB 4; Length 242;  
Best Local Similarity 65.2%; Pred. No. 5.5e-42;  
Matches 103; Conservative 8; Mismatches 27; Indels 20; Gaps 4;

```
Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     56  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 115
```

```

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||||||| | | : | : | | |
Db      116 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 174

Qy      112 ALYLVCGERGFFYTPK-----TRGIVEQCCTSICSLY 143
          : : | | | | | | : : | | | |
Db      175 GIQTLMGVRVAPGVPNPGAPLTLRAVLEKHC---CPLF 209

```

RESULT 5

```

US-09-949-016-8663
; Sequence 8663, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES
THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8663
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8663

```

```

Query Match          58.1%; Score 463; DB 4; Length 242;
Best Local Similarity 65.2%; Pred. No. 5.5e-42;
Matches 103; Conservative 8; Mismatches 27; Indels 20; Gaps 4;

```

```

Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSES IPT 61
          ||||||||||||||||||||||||||||||| | | | | | | | | | | | | | | | | |
Db      56 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSES IPT 115

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||||||| | | : | : | | |
Db      116 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 174

Qy      112 ALYLVCGERGFFYTPK-----TRGIVEQCCTSICSLY 143
          : : | | | | | | : : | | | |
Db      175 GIQTLMGVRVAPGVPNPGAPLTLRAVLEKHC---CPLF 209

```

RESULT 6

```

US-09-949-016-8664
; Sequence 8664, Application US/09949016
; Patent No. 6812339

```



; SEQ ID NO 5  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-284-878-5

Query Match 57.8%; Score 461; DB 3; Length 191;  
Best Local Similarity 70.3%; Pred. No. 6.6e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||| | | : | : | |
Db      61  PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 119

Qy     112  ALYLVCG--ERGFFYTPKTRGIVEQ 134
          : : | | | : | : | |
Db     120  GIQTLMGRLEDG---SPRTGQIFKQ 141
```

RESULT 8

US-09-462-941-1

; Sequence 1, Application US/09462941  
; Patent No. 6608183  
; GENERAL INFORMATION:  
; APPLICANT: Cox III, George N  
; APPLICANT: Bolder Biotechnology, Inc.  
; TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins  
; FILE REFERENCE: 4152-1-PUS  
; CURRENT APPLICATION NUMBER: US/09/462,941  
; CURRENT FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: 60/052,516  
; PRIOR FILING DATE: 1997-07-14  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-462-941-1

Query Match 57.8%; Score 461; DB 4; Length 191;  
Best Local Similarity 70.3%; Pred. No. 6.6e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||| | | : | : | |
Db      61  PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 119

Qy     112  ALYLVCG--ERGFFYTPKTRGIVEQ 134
```

: : | | | :|:| | :|  
Db                 120 GIQTLMGRLDG---SPRTGQIFKQ 141

RESULT 9

US-08-383-621-4

; Sequence 4, Application US/08383621

; Patent No. 5951972

; GENERAL INFORMATION:

; APPLICANT: Daley, Michael J.

; APPLICANT: Buckwalter, Brian L.

; APPLICANT: Cady, Susan M.

; APPLICANT: Shieh, Hong-Ming

; APPLICANT: Bohlen, Peter

; APPLICANT: Seddon, Andrew P.

; TITLE OF INVENTION: Stabilization Of Somatotropins And Other

; TITLE OF INVENTION: Proteins By Modification Of Cysteine Residues

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. Estelle J. Tsevdos

; STREET: 1937 West Main Street, P.O. Box 60

; CITY: Stamford

; STATE: Connecticut

; COUNTRY: U.S.A.

; ZIP: 06904-0060

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/383,621

; FILING DATE: 06-FEB-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/766,142

; FILING DATE: 25-SEP-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Tsevdos, Estelle J.

; REGISTRATION NUMBER: 31,145

; REFERENCE/DOCKET NUMBER: 31,278-01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 203-321-2756

; TELEFAX: 203-321-2971

; TELEX: 203-710-474-4059

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 194 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-383-621-4

Query Match                  57.8%; Score 461; DB 2; Length 194;

Best Local Similarity      70.3%; Pred. No. 6.8e-42;

Matches 102; Conservative   7; Mismatches 20; Indels 16; Gaps 4;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
 |||||  
 Db 4 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 63  
 Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111  
 ||||| : | :  
 Db 64 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 122  
 Qy 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134  
 : : | | : | :  
 Db 123 GIQTLMGRLLEDG---SPRTGQIFKQ 144

RESULT 10

US-08-459-906-4

; Sequence 4, Application US/08459906

; Patent No. 6010999

; GENERAL INFORMATION:

; APPLICANT: Daley, Michael J.

; APPLICANT: Buckwalter, Brian L.

; APPLICANT: Cady, Susan M.

; APPLICANT: Shieh, Hong-Ming

; APPLICANT: Bohlen, Peter

; APPLICANT: Seddon, Andrew P.

; TITLE OF INVENTION: Stabilization of Somatotropins and Other

; TITLE OF INVENTION: Proteins by Modification of Cysteine Residues

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: American Cyanamid Company

; STREET: One Cyanamid Plaza

; CITY: Wayne

; STATE: New Jersey

; COUNTRY: U.S.A.

; ZIP: 07470-8426

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,906

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Darryl L.

; REGISTRATION NUMBER: 34,276

; REFERENCE/DOCKET NUMBER: 31,278-03

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-831-3247

; TELEFAX: 201-831-3305

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 194 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-459-906-4

Query Match 57.8%; Score 461; DB 3; Length 194;  
 Best Local Similarity 70.3%; Pred. No. 6.8e-42;  
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 63

Qy     62 PSNREETQQKSNLELLLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
        ||||||||||||||||||||||||||||| | | : | : | |
Db     64 PSNREETQQKSNLELLLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 122

Qy    112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
        : : | | | : | : | | : |
Db    123 GIQTLMGRLEDG---SPRTGQIFKQ 144
  
```

RESULT 11

US-08-589-028-10

; Sequence 10, Application US/08589028

; Patent No. 6087129

; GENERAL INFORMATION:

; APPLICANT: Newgard, Christopher B.

; APPLICANT: Halban, Philippe

; APPLICANT: No. 6087129mington, Karl D.

; APPLICANT: Clark, Samuel A.

; APPLICANT: Thigpen, Anice E.

; APPLICANT: Quaade, Christian

; APPLICANT: Kruse, Fred

; TITLE OF INVENTION: Recombinant Expression of Proteins From

; TITLE OF INVENTION: Secretory Cell Lines

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: USA

; ZIP: 77210-4433

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/589,028

; FILING DATE: Concurrently Herewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Highlander, Steven L.

; REGISTRATION NUMBER: 47,642

; REFERENCE/DOCKET NUMBER: UTSD:426\HYL

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (512) 418-3000

; TELEFAX: (512) 474-7577

; INFORMATION FOR SEQ ID NO: 10:

```
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 217 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
US-08-589-028-10
```

```
Query Match          57.8%; Score 461; DB 3; Length 217;
Best Local Similarity 70.3%; Pred. No. 7.8e-42;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
```

```
Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
          ||||||||||||||||||||||||||||  | | : | :  | |
Db      87  PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 145

Qy     112  ALYLVCG--ERGFFYTPKTRGIVEQ 134
          : : | | | : | | | : |
Db     146  GIQTLMGRLEDG---SPRTGQIFKQ 167
```

RESULT 12

US-08-784-582-10

```
; Sequence 10, Application US/08784582
; Patent No. 6110707
; GENERAL INFORMATION:
;   APPLICANT: Newgard, Christopher B.
;   APPLICANT: Halban, Philippe A.
;   APPLICANT: No. 6110707mington, Karl D.
;   APPLICANT: Clark, Samuel A.
;   APPLICANT: Thigpen, Anice E.
;   APPLICANT: Quaade, Christian
;   APPLICANT: Kruse, Fred
;   APPLICANT: McGarry, Dennis
;   TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;   TITLE OF INVENTION: SECRETORY CELL LINES
;   NUMBER OF SEQUENCES: 79
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: USA
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/784,582
;     FILING DATE: Concurrently Herewith
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
```



```

; APPLICATION NUMBER: US 60/028,427
; FILING DATE: 15-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-784-582-10

```

```

Query Match          57.8%; Score 461; DB 3; Length 217;
Best Local Similarity 70.3%; Pred. No. 7.8e-42;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

```

Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 86

Qy      62 PSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
        ||||||||||||||||||||||||||||| | | : | : | |
Db      87 PSNREETQQKSNLELLRLISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLEE 145

Qy     112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
        : : | | | : | : | | : |
Db     146 GIQTLMGRLDG---SPRTGQIFKQ 167

```

# RESULT 13

US-08-785-271-10

```

; Sequence 10, Application US/08785271
; Patent No. 6194176
; GENERAL INFORMATION:
; APPLICANT: Newgard, Christopher B.
; APPLICANT: Halban, Philippe A.
; APPLICANT: No. 6194176mington, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Thigpen, Anice E.
; APPLICANT: Quaade, Christian
; APPLICANT: Kruse, Fred
; TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
; TITLE OF INVENTION: SECRETORY CELL LINES
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas

```

```

; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,271
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:513
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-785-271-10

```

```

Query Match          57.8%; Score 461; DB 3; Length 217;
Best Local Similarity 70.3%; Pred. No. 7.8e-42;
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

```

QY      2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

QY      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
        ||||||||||||||||||||||||||||| | | : | : | |
Db      87 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLEE 145

QY     112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
        : : | | | : | : | | : |
Db     146 GIQTLMGRLEDG---SPRTGQIFKQ 167

```

# RESULT 14

US-08-759-628-11

```

; Sequence 11, Application US/08759628
; Patent No. 6225446

```

## GENERAL INFORMATION:

```

; APPLICANT: Altmann, Scott W.
; APPLICANT: Rock, Fernando L.
; APPLICANT: Bazan, J. Fernando
; APPLICANT: Kastelein, Robert A.
; TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMLIAN PROTEINS
; NUMBER OF SEQUENCES: 11

```

```

;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  DNAX Research Institute
;   STREET:    901 California Avenue
;   CITY:      Palo Alto
;   STATE:     California
;   COUNTRY:   USA
;   ZIP:       94304-1104
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:    IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:    PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/759,628
;   FILING DATE:    05-DEC-1996
;   CLASSIFICATION:  435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 60/008,574
;   FILING DATE:    06-DEC-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME:          Ching, Edwin P.
;   REGISTRATION NUMBER:  34,090
;   REFERENCE/DOCKET NUMBER:  DX0552Q
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:    415-852-9196
;   TELEFAX:      415-496-1200
;   INFORMATION FOR SEQ ID NO:  11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:       217 amino acids
;   TYPE:          amino acid
;   STRANDEDNESS:  single
;   TOPOLOGY:     linear
;   MOLECULE TYPE:  protein
;   FEATURE:
;   NAME/KEY:      Peptide
;   LOCATION:       32..53
;   FEATURE:
;   NAME/KEY:      Peptide
;   LOCATION:       94..115
;   FEATURE:
;   NAME/KEY:      Peptide
;   LOCATION:       133..153
;   FEATURE:
;   NAME/KEY:      Peptide
;   LOCATION:       192..210
;   OTHER INFORMATION:  /note= "The peptides above are
;   OTHER INFORMATION:  depicted in Figure 1"
US-08-759-628-11

```

```

Query Match          57.8%;  Score 461;  DB 3;  Length 217;
Best Local Similarity 70.3%;  Pred. No. 7.8e-42;
Matches 102;  Conservative 7;  Mismatches 20;  Indels 16;  Gaps 4;

QY      2  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      |||
Db      27  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

```

```

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
        |||
Db      87 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 145

Qy      112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
        : : | | | :|:| | :|
Db      146 GIQTLMGRL EDG---SPRTGQIFKQ 167

```

RESULT 15

US-09-284-878-1

; Sequence 1, Application US/09284878

; Patent No. 6342375

; GENERAL INFORMATION:

; APPLICANT: Olazaran, Martha Guerrero

; APPLICANT: Saldana, Hugo Barrera

; APPLICANT: Salvado, Jose Maria Viader

; TITLE OF INVENTION: Genetically Modified Methylo trophic P. pastoris Yeast  
for the

; TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone

; FILE REFERENCE: 1829.0010000

; CURRENT APPLICATION NUMBER: US/09/284,878

; CURRENT FILING DATE: 1999-07-21

; PRIOR APPLICATION NUMBER: PCT/MX97/00033

; PRIOR FILING DATE: 1997-10-24

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 217

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-284-878-1

Query Match 57.8%; Score 461; DB 3; Length 217;

Best Local Similarity 70.3%; Pred. No. 7.8e-42;

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
        |||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
        |||
Db      87 PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 145

Qy      112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
        : : | | | :|:| | :|
Db      146 GIQTLMGRL EDG---SPRTGQIFKQ 167

```

Search completed: February 11, 2005, 18:27:07

Job time : 40.5756 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:33 ; Search time 28.5055 Seconds  
(without alignments)  
506.306 Million cell updates/sec

Title: US-10-054-873-7  
Perfect score: 797  
Sequence: 1 MFPTIPLSRLFDNAMLRAHR.....IVEQCCTSICSLYQLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	461	57.8	217	1	STHU somatotropin 1 pre
2	460	57.7	217	2	I67410 somatotropin - rhe
3	426.5	53.5	217	1	STHUV somatotropin 2 pre
4	426.5	53.5	256	1	STHUV2 somatotropin 2 pre
5	407.5	51.1	217	2	I67409 chorionic somatoma
6	405	50.8	217	2	I67411 somatotropin - rhe
7	396	49.7	212	2	I67408 chorionic somatoma
8	396	49.7	217	2	I53267 chorionic somatoma
9	381	47.8	217	1	LCHUC choriomammotropin
10	381	47.8	217	2	E32435 choriomammotropin
11	359.5	45.1	215	2	A26449 choriomammotropin
12	310.5	39.0	216	2	B49159 somatotropin - gol
13	307.5	38.6	190	2	PN0140 somatotropin - sei

14	306.5	38.5	190	1	STHO	somatotropin - hor
15	304.5	38.2	216	1	STMS	somatotropin precu
16	302.5	38.0	216	1	STRT	somatotropin precu
17	302.5	38.0	216	2	S49483	somatotropin precu
18	301.5	37.8	190	2	JK0219	somatotropin - Afr
19	301.5	37.8	216	1	STPG	somatotropin precu
20	301.5	37.8	216	2	I46145	somatotropin precu
21	301.5	37.8	216	2	JC4632	somatotropin precu
22	299.5	37.6	216	2	A37782	somatotropin precu
23	297.5	37.3	190	1	A61584	somatotropin - alp
24	295.5	37.1	190	2	JS0429	somatotropin - Arc
25	289.5	36.3	217	1	STBO	somatotropin precu
26	289.5	36.3	217	1	STGT	somatotropin precu
27	289.5	36.3	217	1	STSH	somatotropin precu
28	289.5	36.3	217	2	S32682	somatotropin - dom
29	278.5	34.9	216	2	JC1514	somatotropin precu
30	277.5	34.8	110	1	INRB	insulin precursor
31	277.5	34.8	110	2	B42179	insulin precursor
32	275.5	34.6	216	2	A60509	somatotropin precu
33	275	34.5	96	2	PC7082	epidermal growth f
34	273.5	34.3	51	1	INEL	insulin - elephant
35	273.5	34.3	51	1	INWHF	insulin - finback
36	273.5	34.3	51	1	INWHP	insulin - sperm wh
37	273.5	34.3	110	2	JQ0178	insulin precursor
38	272	34.1	110	2	A42179	insulin precursor
39	271.5	34.1	51	1	INHY	insulin - hamster
40	270	33.9	110	1	IPHU	insulin precursor
41	268.5	33.7	51	1	INMSSP	insulin - Egyptian
42	268.5	33.7	191	2	A60625	somatotropin - gre
43	267.5	33.6	51	2	A59151	insulin precursor
44	266.5	33.4	105	1	IPBO	insulin precursor
45	265.5	33.3	110	2	I48166	insulin precursor

#### ALIGNMENTS

##### RESULT 1

##### STHU

somatotropin 1 precursor [validated] - human

N;Alternate names: growth hormone 1; hGH-N; pituitary somatotropin

N;Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, short form

C;Species: Homo sapiens (man)

C;Date: 24-Apr-1984 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004

C;Accession: A93731; A32435; A93694; A94247; A90051; A93397; A93778; A91764; A90217; A92311; A61466; S09685; I84549; A01510

R;DeNoto, F.M.; Moore, D.D.; Goodman, H.M.

Nucleic Acids Res. 9, 3719-3730, 1981

A;Title: Human growth hormone DNA sequence and mRNA structure: possible alternative splicing.

A;Reference number: A93731; MUID:82014939; PMID:6269091

A;Accession: A93731

A;Molecule type: DNA

A;Residues: 1-217 <DEN>

A;Cross-references: UNIPROT:P01241; GB:V00520

A;Note: the 20K short form somatotropin lacks residues 58-72 (32-46 in the active hormone) as the result of splicing at the alternate junction of the second intron during mRNA processing  
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg, P.H.  
Genomics 4, 479-497, 1989  
A;Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.  
A;Reference number: A32435; MUID:89307277; PMID:2744760  
A;Accession: A32435  
A;Molecule type: DNA  
A;Residues: 1-217 <CHE>  
A;Cross-references: GB:J03071; NID:g183148; PIDN:AAA52549.1; PID:g183149  
R;Roskam, W.; Rougeon, F.  
Nucleic Acids Res. 7, 305-320, 1979  
A;Title: Molecular cloning and nucleotide sequence of the human growth hormone structural gene.  
A;Reference number: A93694; MUID:80034477; PMID:386281  
A;Accession: A93694  
A;Molecule type: mRNA  
A;Residues: 1-217 <ROS>  
A;Cross-references: GB:V00519  
A;Note: 35-Pro was also found  
R;Martial, J.A.; Hallewell, R.A.; Baxter, J.D.; Goodman, H.M.  
Science 205, 602-607, 1979  
A;Title: Human growth hormone: complementary DNA cloning and expression in bacteria.  
A;Reference number: A94247; MUID:79203293; PMID:377496  
A;Accession: A94247  
A;Molecule type: mRNA  
A;Residues: 1-217 <MAR>  
R;Li, C.H.; Dixon, J.S.; Liu, W.K.  
Arch. Biochem. Biophys. 133, 70-91, 1969  
A;Title: Human pituitary growth hormone. XIX. The primary structure of the hormone.  
A;Reference number: A90048; MUID:69289202; PMID:5810834  
A;Contents: annotation  
R;Li, C.H.; Dixon, J.S.  
Arch. Biochem. Biophys. 146, 233-236, 1971  
A;Title: Human pituitary growth hormone. XXXII. The primary structure of the hormone: revision.  
A;Reference number: A90051; MUID:72143935; PMID:5144027  
A;Accession: A90051  
A;Molecule type: protein  
A;Residues: 27-94;96-217 <LIC>  
R;Niall, H.D.  
Nature New Biol. 230, 90-91, 1971  
A;Title: Revised primary structure for human growth hormone.  
A;Reference number: A93397; MUID:71139765; PMID:5279046  
A;Accession: A93397  
A;Molecule type: protein  
A;Residues: 27-51 <NIA>  
R;Niall, H.D.; Hogan, M.L.; Sauer, R.; Rosenblum, I.Y.; Greenwood, F.C.  
Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971  
A;Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution from a primordial peptide by gene reduplication.  
A;Reference number: A93778; MUID:71153968; PMID:5279528

A;Accession: A93778  
 A;Molecule type: protein  
 A;Residues: 119-120;157-159 <NI2>  
 R;Niall, H.D.  
 in Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop Prolactin,  
 Griffiths, K., ed., pp.13-20, Alpha Omega Alpha Press, Cardiff, Wales, 1972  
 A;Title: The chemistry of the human lactogenic hormones.  
 A;Reference number: A94427  
 A;Contents: annotation; somatotropin revision  
 R;Bewley, T.A.; Dixon, J.S.; Li, C.H.  
 Int. J. Pept. Protein Res. 4, 281-287, 1972  
 A;Title: Sequence comparison of human pituitary growth hormone, human chorionic  
 somatomammotropin, and ovine pituitary growth and lactogenic hormones.  
 A;Reference number: A91764; MUID:73092028; PMID:4675454  
 A;Accession: A91764  
 A;Molecule type: protein  
 A;Residues: 27-217 <BEW>  
 R;Lewis, U.J.; Bonewald, L.F.; Lewis, L.J.  
 Biochem. Biophys. Res. Commun. 92, 511-516, 1980  
 A;Title: The 20,000-dalton variant of human growth hormone: location of the  
 amino acid deletions.  
 A;Reference number: A90217; MUID:80130196; PMID:7356479  
 A;Contents: somatotropin, 20K short variant  
 A;Accession: A90217  
 A;Molecule type: protein  
 A;Residues: 46-57;73-80 <LEW>  
 R;Chapman, G.E.; Rogers, K.M.; Brittain, T.; Bradshaw, R.A.; Bates, O.J.;  
 Turner, C.; Cary, P.D.; Crane-Robinson, C.  
 J. Biol. Chem. 256, 2395-2401, 1981  
 A;Title: The 20,000 molecular weight variant of human growth hormone.  
 Preparation and some physical and chemical properties.  
 A;Reference number: A92311; MUID:81117361; PMID:7462247  
 A;Contents: somatotropin, 20K short variant  
 A;Accession: A92311  
 A;Molecule type: protein  
 A;Residues: 27-57;73-79 <CHA>  
 R;Singh, R.N.P.; Seavey, B.K.; Lewis, L.J.; Lewis, U.J.  
 J. Protein Chem. 2, 425-436, 1983  
 A;Title: Human growth hormone peptide 1-43: isolation from pituitary glands.  
 A;Reference number: A61466  
 A;Accession: A61466  
 A;Molecule type: protein  
 A;Residues: 27-69 <SIN>  
 A;Note: growth hormone 5K peptide has insulin potentiating activity; its  
 physiological production is uncertain  
 R;Robson, V.M.J.; Rae, I.D.; NG, F.  
 Biol. Chem. Hoppe-Seyler 371, 423-431, 1990  
 A;Title: Identification of the aspartimide structure in a previously-reported  
 peptide.  
 A;Reference number: S09685; MUID:90334745; PMID:2378679  
 A;Accession: S09685  
 A;Molecule type: protein  
 A;Residues: 27-34,'L',36-47 <ROB>  
 R;de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.  
 Science 255, 306-312, 1992  
 A;Title: Human growth hormone and extracellular domain of its receptor: crystal  
 structure of the complex.



A;Reference number: A41728; MUID:92196577; PMID:1549776  
A;Contents: annotation; X-ray crystallography, 2.8 angstroms  
A;Note: the structure of the complex with growth hormone receptor is described  
R;Gray, G.L.; Baldridge, J.S.; McKeown, K.S.; Heyneker, H.L.; Chang, C.N.  
Gene 39, 247-254, 1985  
A;Title: Periplasmic production of correctly processed human growth hormone in  
Escherichia coli: natural and bacterial signal sequences are interchangeable.  
A;Reference number: I41126; MUID:86137393; PMID:3912261  
A;Accession: I84549  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-26 <RES>  
A;Cross-references: GB:M14398; NID:g183158; PIDN:AAA52554.1; PID:g183159  
C;Comment: The gene for this hormone is transcribed only in somatotrophic cells  
of the anterior pituitary.  
C;Comment: About 90% of somatotropin is the 22K long form.  
C;Genetics:  
A;Gene: GDB:GH1  
A;Cross-references: GDB:119982; OMIM:139250  
A;Map position: 17q23.1-17q23.3  
A;Introns: 4/1; 57/3; 97/3; 152/3  
C;Superfamily: prolactin  
C;Keywords: alternative splicing; hormone; pituitary  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-217/Product: somatotropin 1, long form #status experimental <SOL>  
F;27-69/Product: growth hormone 5K peptide #status experimental <5KP>  
F;27-57,73-217/Product: somatotropin 1, short form #status experimental <SOS>  
F;79-191,208-215/Disulfide bonds: #status experimental

Query Match 57.8%; Score 461; DB 1; Length 217;  
Best Local Similarity 70.3%; Pred. No. 2.8e-37;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      |||
Db      27  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 86

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLVE 111
      |||
Db      87  PSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 145

Qy      112 ALYLVCG--ERGFFYTPKTRGIVEQ 134
      : : | | | :|:| | :|
Db      146 GIQTLMGRLEDG---SPRTGQIFKQ 167

```

## RESULT 2

I67410

somatotropin - rhesus macaque

N;Alternate names: growth hormone

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C;Accession: I67410; A05094

R;Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993

A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.

A;Reference number: I53267; MUID:94008724; PMID:8404617

A;Accession: I67410

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-217 <RES>

A;Cross-references: UNIPROT:P33093; GB:L16556; NID:g293114; PIDN:AAA18842.1; PID:g293115

R;Li, C.H.; Chung, D.; Lahm, H.W.; Stein, S.

Arch. Biochem. Biophys. 245, 287-291, 1986

A;Title: The primary structure of monkey pituitary growth hormone.

A;Reference number: A05094; MUID:86129460; PMID:3080959

A;Accession: A05094

A;Molecule type: protein

A;Residues: 27-99,'Q',101-178,'D',180-217 <LIC>

A;Note: the monkey species is not identified in the reference

R;Raben, M.S.

Science 125, 883-884, 1957

A;Title: Preparation of growth hormone from pituitaries of man and monkey.

A;Reference number: A44774

A;Contents: annotation; identification of source organism

C;Superfamily: prolactin

Query Match 57.7%; Score 460; DB 2; Length 217;  
Best Local Similarity 98.9%; Pred. No. 3.5e-37;  
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
|||||

Db 27 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92  
|||||

Db 87 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 117

### RESULT 3

STHUV

somatotropin 2 precursor - human

N;Alternate names: growth hormone 2; growth hormone variant; hGH-V; placental somatotropin

N;Contains: somatotropin 2, long splice form; somatotropin 2, short splice form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1982 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004

C;Accession: D32435; B28072; A01511; I52104; A60711

R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg, P.H.

Genomics 4, 479-497, 1989

A;Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.

A;Reference number: A32435; MUID:89307277; PMID:2744760

A;Accession: D32435

A;Molecule type: DNA

A;Residues: 1-217 <CHE>

A;Cross-references: UNIPROT:P01242; GB:J03071; NID:g183148; PIDN:AAA52552.1; PID:g183152  
 R;Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhaber, S.A.  
 J. Biol. Chem. 263, 9001-9006, 1988  
 A;Title: Two distinct species of human growth hormone-variant mRNA in the human placenta predict the expression of novel growth hormone proteins.  
 A;Reference number: A92725; MUID:88243769; PMID:3379057  
 A;Accession: B28072  
 A;Molecule type: mRNA  
 A;Residues: 1-217 <COO>  
 R;Seeburg, P.H.  
 DNA 1, 239-249, 1982  
 A;Title: The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone.  
 A;Reference number: A01511; MUID:83182010; PMID:7169009  
 A;Accession: A01511  
 A;Molecule type: DNA  
 A;Residues: 1-34,'P',36-217 <SEE>  
 R;Igout, A.; Scippo, M.L.; Frankenne, F.; Hennen, G.  
 Arch. Int. Physiol. Biochim. 96, 63-67, 1988  
 A;Title: Cloning and nucleotide sequence of placental hGH-V cDNA.  
 A;Reference number: I52104; MUID:89024984; PMID:2460050  
 A;Accession: I52104  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-217 <IGO>  
 A;Cross-references: GB:M38451; NID:g183179; PIDN:AAA35891.1; PID:g183180  
 R;Frankenne, F.; Scippo, M.L.; Van Beeumen, J.; Igout, A.; Hennen, G.  
 J. Clin. Endocrinol. Metab. 71, 15-18, 1990  
 A;Title: Identification of placental human growth hormone as the growth hormone-V gene expression product.  
 A;Reference number: A60711; MUID:90317018; PMID:2196278  
 A;Accession: A60711  
 A;Molecule type: protein  
 A;Residues: 27-44;46-57 <FRA>  
 A;Experimental source: tissue placenta  
 A;Note: partial glycosylation was demonstrated by lectin binding  
 C;Comment: This gene is expressed by the placenta.  
 C;Genetics:  
 A;Gene: GDB:GH2  
 A;Cross-references: GDB:119983; OMIM:139240  
 A;Map position: 17q22-17q24  
 A;Introns: 4/1; 57/3; 97/3; 152/3  
 C;Superfamily: prolactin  
 C;Keywords: alternative splicing; glycoprotein; hormone; placenta  
 F;1-26/Domain: signal sequence #status predicted <SIG>  
 F;27-217/Product: somatotropin 2, long splice form #status predicted <SOL>  
 F;27-57,73-217/Product: somatotropin 2, short splice form #status predicted <SOS>  
 F;79-191,208-215/Disulfide bonds: #status predicted  
 F;166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5%; Score 426.5; DB 1; Length 217;  
 Best Local Similarity 78.1%; Pred. No. 6.3e-34;  
 Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61

```

          |||:|:|
Db      27 FPTIPLSRLFDNAMLRRRLYQLAYDTYQEFEEAYILKEQKYSFLQNPQTSLCFSESIPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQL-----GTGPRFVNQHL 104
          |||:|:|
Db      87 PSNRVKTQQKSNLELLRISLLLIQSWLEPVQLLRVVFANSLVYGASDSNVYRHL 140

```

#### RESULT 4

STHUV2

somatotropin 2 precursor, splice form 2 - human

N;Alternate names: growth hormone variant-2; placental somatotropin form 2

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1989 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004

C;Accession: A28072

R;Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhaver, S.A.

J. Biol. Chem. 263, 9001-9006, 1988

A;Title: Two distinct species of human growth hormone-variant mRNA in the human placenta predict the expression of novel growth hormone proteins.

A;Reference number: A92725; MUID:88243769; PMID:3379057

A;Accession: A28072

A;Molecule type: mRNA

A;Residues: 1-256 <COO>

A;Cross-references: UNIPROT:P01242

A;Note: an alternative splice junction for intron 4 is used

C;Genetics:

A;Gene: GDB:GH2

A;Cross-references: GDB:119983; OMIM:139240

A;Map position: 17q22-17q24

A;Introns: 4/1; 57/3; 97/3; 152/3

C;Superfamily: prolactin

C;Keywords: alternative splicing; hormone; placenta

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-256/Product: somatotropin 2 splice form 2 #status predicted <MAT>

```

Query Match          53.5%;   Score 426.5;   DB 1;   Length 256;
Best Local Similarity 78.1%;   Pred. No. 7.6e-34;
Matches   89;   Conservative    4;   Mismatches   10;   Indels    11;   Gaps      1;

```

```

Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
          |||:|:|
Db      27 FPTIPLSRLFDNAMLRRRLYQLAYDTYQEFEEAYILKEQKYSFLQNPQTSLCFSESIPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQL-----GTGPRFVNQHL 104
          |||:|:|
Db      87 PSNRVKTQQKSNLELLRISLLLIQSWLEPVQLLRVVFANSLVYGASDSNVYRHL 140

```

#### RESULT 5

I67409

chorionic somatomammotropin-3 - rhesus macaque

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C;Accession: I67409

R;Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993



```

Qy      115 LVCGERG 121
          |   | |
Db      140 LKKLEEG 146

```

## RESULT 7

I67408

chorionic somatomammotropin-2 - rhesus macaque (fragment)

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 31-May-1996 #sequence revision 31-May-1996 #text change 09-Jul-2004

C;Accession: I67408

R;Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993

A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.

A;Reference number: I53267; MUID:94008724; PMID:8404617

A;Accession: I67408

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-212 <RES>

A;Cross-references: UNIPROT:Q07368; GB:L16553; NID:g293110; PIDN:AAA18840.1;  
PID:g293111

C;Superfamily: prolactin

Query Match 49.7%; Score 396; DB 2; Length 212;  
Best Local Similarity 82.2%; Pred. No. 5.5e-31;  
Matches 74; Conservative 11; Mismatches 5; Indels 0; Gaps 0;

QY            3 PTIPLSRLFDNAMLRAHRLHQAFDITYQEFEAYIPKEQKYSFLQNPQTSLSFSESIPTP 62  
|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

Db 23 PSVPLSR LFDHAMIOAHRLHOLAFDTYOEFE EAYIPKEKKHSLMENPOASFCFADSIPTP 82

Qv 63 SNREETOOKSNLELLRISLLLIOSWLEPVO 92

\_\_\_\_\_

Db 83 SNLEETOOKSNLELLRISLLLIOSWLEPVO 112

## RESULT 8

T53267

chorionic somatomammotropin-1 - rhesus macaque

C;Species: Macaca mulatta (rhesus macaque)

C;Date: 31-May-1996 #sequence revision 31-May-1996 #text change 09-Jul-2004

C;Accession: I53267

R;Golos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993

A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.

A;Reference number: I53267; MUID:94008724; PMID:8404617

A;Accession: I53267

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-217 <RES>

A/Cross-references: UNIPROT:Q07367; GB:L16552; NID:g293108; PIDN:AAA18839.1;  
PID:g293109

```
Query Match          49.7%; Score 396; DB 2; Length 217;  
Best Local Similarity 82.2%; Pred. No. 5.7e-31;  
Matches    74; Conservative   11; Mismatches     5; Indels      0; Gaps      0;  
  
Qy           3 PTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPTSLSFSESIPTP 62  
              |:|||||::||:|||||:|||||:|||||:|:| ::||| | |:|||||  
Db           28 PSVPLSRLFDHAMIQAHLRHQLAFDITYQEFEEAYIPKEKKHSMLMENPQASFCFADSIPTP 87  
  
Qy           63 SNREETQQKSNELELLRISLLLIQSWLEPVQ 92  
              || |||||:|||||:|||||:|||||:|||||  
Db           88 SNLEETOOKSNELELLRISLLLIQSWLEPVQ 117
```

## LCHUC

chorionmammotropin A precursor [validated] - human  
N;Alternate names: chorionic somatomammotropin 1; placental lactogen  
C;Species: Homo sapiens (man)  
C;Date: 23-Oct-1981 #sequence\_revision 23-Oct-1981 #text\_change 09-Jul-2004  
C;Accession: C32435; A94422; I52342; A93833; A93192; A90054; A94427; A61283;  
I55229; I59658; A01512  
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.;  
Seeburg, P.H.  
Genomics 4, 479-497, 1989  
A;Title: The human growth hormone locus: nucleotide sequence, biology, and  
evolution.  
A;Reference number: A32435; MUID:89307277; PMID:2744760  
A;Accession: C32435  
A;Molecule type: DNA  
A;Residues: 1-217 <CHE>  
A;Cross-references: UNIPROT:P01243; GB:J03071; NID:g183148; PIDN:AAA52551.1;  
PID:g183151  
R;Goodman, H.M.; DeNoto, F.; Fiddes, J.C.; Hallelwell, R.A.; Page, G.S.; Smith,  
S.; Tischer, E.  
in Mobilization and Reassembly of Genetic Information, Scott, W.A., Werner, R.,  
Joseph, D.R., and Schultz, J., eds., pp.155-179, Academic Press, New York, 1980  
A;Reference number: A94422  
A;Accession: A94422  
A;Molecule type: mRNA  
A;Residues: 1-217 <GOO>  
R;Tanaka, M.; Masuda, N.; Watahiki, M.; Yamakawa, M.; Shimizu, K.; Nagai, J.;  
Nakashima, K.  
Biochem. Int. 16, 287-292, 1988  
A;Title: cDNA cloning of human chorionic somatomammotropin-1 mRNA whose  
transcription was initiated at the 5' region of the TATA box.  
A;Reference number: I52342; MUID:88209096; PMID:2835050  
A;Accession: I52342  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-3 <TAN>  
A;Cross-references: GB:M35419; NID:g506822  
R;Sherwood, L.M.; Burstein, Y.; Schechter, I.  
Proc. Natl. Acad. Sci. U.S.A. 76, 3819-3823, 1979  
A;Title: Primary structure of the NH-2-terminal extra piece of the precursor to  
human placental lactogen.

A;Reference number: A93833; MUID:80034970; PMID:291043  
 A;Accession: A93833  
 A;Molecule type: protein  
 A;Residues: 1,3-26 <SHE>  
 A;Experimental source: placenta  
 R;Shine, J.; Seeburg, P.H.; Martial, J.A.; Baxter, J.D.; Goodman, H.M.  
 Nature 270, 494-499, 1977  
 A;Title: Construction and analysis of recombinant DNA for human chorionic somatomammotropin.  
 A;Reference number: A93192; MUID:78071761; PMID:593368  
 A;Accession: A93192  
 A;Molecule type: DNA  
 A;Residues: 50-217 <SHI>  
 A;Experimental source: placenta  
 R;Li, C.H.; Dixon, J.S.; Chung, D.  
 Arch. Biochem. Biophys. 155, 95-110, 1973  
 A;Title: Amino acid sequence of human chorionic somatomammotropin.  
 A;Reference number: A90054; MUID:73201971; PMID:4712450  
 A;Accession: A90054  
 A;Molecule type: protein  
 A;Residues: 27-217 <LIC>  
 A;Experimental source: placenta  
 R;Niall, H.D.  
 in Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop Prolactin,  
 Griffiths, K., ed., pp.13-20, Alpha Omega Alpha Press, Cardiff, Wales, 1972  
 A;Title: The chemistry of the human lactogenic hormones.  
 A;Reference number: A94427  
 A;Accession: A94427  
 A;Molecule type: protein  
 A;Residues: 27-217 <NIA>  
 A;Experimental source: placenta  
 R;Nic A Bhaird, N.; Tipton, K.F.  
 Biochem. Soc. Trans. 19, 20S, 1991  
 A;Title: Catechol-O-methyltransferase from human placenta: purification and some properties.  
 A;Reference number: A61283; MUID:91244006; PMID:2037148  
 A;Accession: A61283  
 A;Molecule type: protein  
 A;Residues: 27-46 <NIC>  
 A;Note: choriomammotropin apparently copurified with placental catechol-O-methyltransferase  
 R;Sherwood, L.M.; Handwerger, S.; McLaurin, W.D.; Lanner, M.  
 Nature New Biol. 233, 59-61, 1971  
 A;Title: Amino-acid sequence of human placental lactogen.  
 A;Reference number: A93401; MUID:72016313; PMID:5286363  
 A;Contents: annotation  
 R;Sherwood, L.M.; Handwerger, S.; McLaurin, W.D.; Lanner, M.  
 Nature New Biol. 235, 64, 1972  
 A;Reference number: A93405  
 A;Contents: annotation  
 R;Schneider, A.B.; Kowalski, K.; Russell, J.; Sherwood, L.M.  
 J. Biol. Chem. 254, 3782-3787, 1979  
 A;Title: Identification of the interchain disulfide bonds of dimeric human placental lactogen.  
 A;Reference number: A92251; MUID:79173081; PMID:438159  
 A;Contents: annotation; dimeric disulfide bonds  
 R;Selby, M.J.; Barta, A.; Baxter, J.D.; Bell, G.I.; Eberhardt, N.L.



J. Biol. Chem. 259, 13131-13138, 1984  
 A;Title: Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional promoter elements.  
 A;Reference number: I55229; MUID:85030426; PMID:6208192  
 A;Accession: I55229  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-217 <RES>  
 A;Cross-references: GB:K02401; NID:gl81120; PIDN:AAA52115.1; PID:gl81121  
 R;Seeburg, P.H.; Shine, J.; Martial, J.A.; Ullrich, A.; Goodman, H.  
 Trans. Assoc. Am. Physicians 90, 109-116, 1977  
 A;Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.  
 A;Reference number: I59658; MUID:78160787; PMID:611657  
 A;Accession: I59658  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 160-217 <RE2>  
 A;Cross-references: GB:M25118; NID:gl81124; PIDN:AAA35721.1; PID:gl81125  
 C;Genetics:  
 A;Gene: GDB:CSH1  
 A;Cross-references: GDB:119084; OMIM:150200  
 A;Map position: 17q22-17q24  
 A;Introns: 4/1; 57/3; 97/3; 152/3  
 C;Superfamily: prolactin  
 C;Keywords: hormone; placenta  
 F;1-26/Domain: signal sequence #status experimental <SIG>  
 F;27-217/Product: choriomammotropin A #status experimental <MAT>  
 F;79-191/Disulfide bonds: #status experimental  
 F;208-215/Disulfide bonds: (in monomeric form) #status experimental  
 F;208/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental  
 F;215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental

Query Match 47.8%; Score 381; DB 1; Length 217;  
 Best Local Similarity 82.0%; Pred. No. 1.6e-29;  
 Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

```
Qy      4 TIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPTPS 63
      |:|||||:||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      29 TVPLSRLFDHAMLQAHRAHQLAIDTYQEFEEYIPKDQKYSFLHDSQTSFCFSDSIPTPS 88

Qy      64 NREETQQKSNLELLLRISLLLIQSWLEPVQ 92
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      89 NMEETQQKSNLELLLRISLLLIQSWLEPVR 117
```

RESULT 10  
 E32435  
 choriomammotropin B precursor - human  
 N;Alternate names: chorionic somatomammotropin 2  
 C;Species: Homo sapiens (man)  
 C;Date: 29-Dec-1989 #sequence\_revision 29-Dec-1989 #text\_change 09-Jul-2004  
 C;Accession: E32435  
 R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.;  
 Seeburg, P.H.  
 Genomics 4, 479-497, 1989  
 A;Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.

A;Reference number: A32435; MUID:89307277; PMID:2744760  
 A;Accession: E32435  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-217 <CHE>  
 A;Cross-references: UNIPROT:Q14407; GB:J03071; NID:g183148; PIDN:AAA52553.1;  
 PID:g183153  
 C;Genetics:  
 A;Gene: GDB:CSH2  
 A;Cross-references: GDB:119813; OMIM:118820  
 A;Map position: 17q22-17q24  
 C;Superfamily: prolactin

Query Match 47.8%; Score 381; DB 2; Length 217;  
 Best Local Similarity 82.0%; Pred. No. 1.6e-29;  
 Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

```
Qy      4 TIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPTS 63
          |:|||||||:||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      29 TVPLSRLFDHAMLQAHRAHQLAIDTYQFEETYIPKDQKYSFLHDSQTSFCFSDSIPTS 88

Qy      64 NREETQQKSNLELLRLISLLLIQSWLEPVQ 92
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      89 NMEETQQKSNLELLRLISLLLIQSWLEPVR 117
```

RESULT 11

A26449

choriomammotropin precursor (allele hCS-3) - human

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004

C;Accession: A26449

R;Hirt, H.; Kimelman, J.; Birnbaum, M.J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.; Barta, A.

DNA 6, 59-70, 1987

A;Title: The human growth hormone gene locus: structure, evolution, and allelic variations.

A;Reference number: A26449; MUID:87161235; PMID:3030680

A;Accession: A26449

A;Molecule type: DNA

A;Residues: 1-215 <HIR>

A;Cross-references: UNIPROT:P01243

C;Superfamily: prolactin

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-215/Product: choriomammotropin, hCS-3 allele #status predicted <MAT>

Query Match 45.1%; Score 359.5; DB 2; Length 215;  
 Best Local Similarity 80.5%; Pred. No. 1.9e-27;  
 Matches 70; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

```
Qy      4 TIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPTS 63
          |:|||||||:||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      29 TVPLSRLFDHAMLQAHRAHQLAIDTYQFEETYIPKDQKYSFLHDSQTSFCFSDSIPTS 88

Qy      64 NREETQQKSNLELLRLISLLLIQSWLEP 90
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      89 NMEETQQKSNLELLRL-LLLIQSWLEP 114
```

# RESULT 12

B49159

somatotropin - golden hamster

N;Alternate names: growth hormone

C;Species: *Mesocricetus auratus* (golden hamster)

C;Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: B49159

R;Southard, J.N.; Sanchez-Jimenez, F.; Campbell, G.T.; Talamantes, F.  
Endocrinology 129, 2965-2971, 1991

A;Title: Sequence and expression of hamster prolactin and growth hormone messenger RNAs.

A;Reference number: A49159; MUID:92063850; PMID:1954881

A;Accession: B49159

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-216 <SOU>

A;Cross-references: UNIPROT:P37886; GB:S66299; NID:g239355; PIDN:AAB20368.1;  
PID:g239356

A;Note: sequence extracted from NCBI backbone (NCBIN:66299, NCBIP:66300)

C;Superfamily: prolactin

Query Match 39.0%; Score 310.5; DB 2; Length 216;

Best Local Similarity 67.0%; Pred. No. 1.1e-22;

Matches 61; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
|| :||| || ||:||| ||||| |||:||| ||||: |:|| :|| ||: |||:||

Db 27 FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYIS-IQNAQTAFCFSETIPA 85

Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92  
|: :|| ||:|::||| ||||| |||

Db 86 PTGKEEAQQRSDMELLRFSLLLIQSWLGPVQ 116

# RESULT 13

PN0140

somatotropin - sei whale

N;Alternate names: growth hormone

C;Species: *Balaenoptera borealis* (sei whale)

C;Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 09-Jul-2004

C;Accession: PN0140

R;Yudaev, N.A.; Pankov, Y.A.; Bulatov, A.A.; Osipova, T.A.  
Biokhimiia 47, 1059-1069, 1982

A;Title: Amino acid sequence of seiwhale somatotropin.

A;Reference number: PN0140; MUID:83000569; PMID:7115813

A;Accession: PN0140

A;Molecule type: protein

A;Residues: 1-190 <YUD>

A;Cross-references: UNIPROT:P33092

A;Note: article in Russian with English abstract

C;Superfamily: prolactin

C;Keywords: growth factor; hormone

F;52-163,180-188/Disulfide bonds: #status predicted

Query Match 38.6%; Score 307.5; DB 2; Length 190;

Best Local Similarity 67.0%; Pred. No. 1.8e-22;  
Matches 61; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

```
Qy      2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      || :||| || ||:||| ||:|| |||:||| ||||: |:| |||| |:: ||| |||
Db      1 FPAMPLSSLFANAVLRAQHLLHELAADTYKEFERAYIPEGQRY-FLQNAQSTGCFSEVIPT 59

Qy      62 PSNREETQQKSNLELLLRISLLLIQSWLEPVQ 92
      |:|::| ||:|::||| ||||| |||
Db      60 PANKDEAQQRSDVELLRFSLLLIQSWLGPVQ 90
```

#### RESULT 14

STHO

somatotropin - horse

N;Alternate names: growth hormone

C;Species: Equus caballus (domestic horse)

C;Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 23-Aug-1996

C;Accession: A91772; A91395; A91383; A90240; A01514

R;Zakin, M.M.; Poskus, E.; Langton, A.A.; Ferrara, P.; Santome, J.A.; Dellacha, J.M.; Paladini, A.C.

Int. J. Pept. Protein Res. 8, 435-444, 1976

A;Title: Primary structure of equine growth hormone.

A;Reference number: A91772; MUID:77005410; PMID:965151

A;Accession: A91772

A;Molecule type: protein

A;Residues: 1-190 <ZAK>

R;Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.

FEBS Lett. 34, 353-355, 1973

A;Title: The amino acid sequence of equine growth hormone.

A;Reference number: A91395; MUID:74020362; PMID:4747849

A;Accession: A91395

A;Molecule type: protein

A;Residues: 1-190 <ZA2>

R;Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.

FEBS Lett. 25, 77-82, 1972

A;Title: Amino acid sequences around the cystine residues in equine growth hormone.

A;Reference number: A91383

A;Accession: A91383

A;Molecule type: protein

A;Residues: 42-69;157-190 <ZA3>

R;Oliver, L.; Hartree, A.S.

Biochem. J. 109, 19-24, 1968

A;Title: Amino acid sequences around the cystine residues in horse growth hormone.

A;Reference number: A90240; MUID:68368390; PMID:4876100

A;Accession: A90240

A;Molecule type: protein

A;Residues: 176-190 <OLI>

C;Superfamily: prolactin

C;Keywords: hormone; pituitary

F;52-163,180-188/Disulfide bonds: #status experimental

Query Match 38.5%; Score 306.5; DB 1; Length 190;

Best Local Similarity 65.2%; Pred. No. 2.3e-22;

Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      || :||| || ||:||| ||||| |||:||| ||||: |:|| :|| | : |||:||
Db      1  FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYIS-IQNAQA AFCFSETIPA 59

Qy      62 PSNREETQQKSNLELLLRISLLLIQSWLEPVQL 93
      |: ::| ||:|::||| ||||| ||||| ||||
Db      60 PTGKDEAQQRSDMELLRFSLLLIQSWLGPVQL 91

```

RESULT 15

STMS

somatotropin precursor - mouse

N;Alternate names: growth hormone

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004

C;Accession: B23911

R;Linzer, D.I.H.; Talamantes, F.

J. Biol. Chem. 260, 9574-9579, 1985

A;Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression of these mRNAs during pregnancy.

A;Reference number: A92548; MUID:85261358; PMID:2991252

A;Accession: B23911

A;Molecule type: mRNA

A;Residues: 1-216 <LIN>

A;Cross-references: UNIPROT:P06880; GB:X02891; GB:K03232; NID:g51067;

PIDN:CAA26650.1; PID:g51068

C;Superfamily: prolactin

C;Keywords: anterior pituitary; growth factor; hormone

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-216/Product: somatotropin #status predicted <STN>

F;78-189,206-214/Disulfide bonds: #status predicted

Query Match 38.2%; Score 304.5; DB 1; Length 216;

Best Local Similarity 64.8%; Pred. No. 4.2e-22;

Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      || :||| || ||:||| ||||| |||:||| ||||: |:|| :|| | : |||:||
Db      27 FPAMPLSSLFSNAVLRAQHLHQLAADTYKEFERAYIPEGQRYIS-IQNAQA AFCFSETIPA 85

Qy      62 PSNREETQQKSNLELLLRISLLLIQSWLEPVQ 92
      |: :|| ||:|::||| ||||| ||||| ||||
Db      86 PTGKEEAQQRTDMELLRFSLLLIQSWLGPVQ 116

```

Search completed: February 11, 2005, 18:24:36

Job time : 28.5055 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 18:23:02 ; Search time 109.041 Seconds  
(without alignments)  
449.487 Million cell updates/sec

Title: US-10-054-873-7  
Perfect score: 797  
Sequence: 1 MFPTIPLSRLFDNAMLRAHR.....IVEQCCTSICSLYQLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1376875 seqs, 326749119 residues

Total number of hits satisfying chosen parameters: 1376875

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	797	100.0	150	13	US-10-054-873-7	Sequence 7, Appli
2	555.5	69.7	107	13	US-10-054-873-6	Sequence 6, Appli
3	470	59.0	92	13	US-10-054-873-2	Sequence 2, Appli
4	470	59.0	134	10	US-09-819-094-24	Sequence 24, Appl
5	470	59.0	134	15	US-10-714-067-24	Sequence 24, Appl
6	466	58.5	188	15	US-10-621-693-18	Sequence 18, Appl
7	466	58.5	192	10	US-09-819-094-23	Sequence 23, Appl
8	466	58.5	192	15	US-10-621-693-8	Sequence 8, Appli
9	466	58.5	192	15	US-10-621-693-78	Sequence 78, Appl
10	466	58.5	192	15	US-10-621-693-86	Sequence 86, Appl
11	466	58.5	192	15	US-10-714-067-23	Sequence 23, Appl
12	466	58.5	193	15	US-10-621-693-42	Sequence 42, Appl
13	466	58.5	206	15	US-10-621-693-72	Sequence 72, Appl
14	466	58.5	391	15	US-10-621-693-51	Sequence 51, Appl
15	466	58.5	574	15	US-10-621-693-32	Sequence 32, Appl
16	466	58.5	576	15	US-10-621-693-39	Sequence 39, Appl
17	466	58.5	589	15	US-10-621-693-53	Sequence 53, Appl
18	466	58.5	786	15	US-10-621-693-55	Sequence 55, Appl
19	466	58.5	810	15	US-10-621-693-76	Sequence 76, Appl
20	464	58.2	191	16	US-10-658-834A-875	Sequence 875, App
21	463	58.1	191	16	US-10-658-834A-866	Sequence 866, App
22	463	58.1	191	16	US-10-658-834A-876	Sequence 876, App
23	463	58.1	191	16	US-10-658-834A-887	Sequence 887, App
24	463	58.1	213	17	US-10-741-600-954	Sequence 954, App
25	462	58.0	191	16	US-10-658-834A-867	Sequence 867, App
26	462	58.0	191	16	US-10-658-834A-881	Sequence 881, App
27	462	58.0	191	16	US-10-658-834A-888	Sequence 888, App
28	461	57.8	191	10	US-09-984-010-23	Sequence 23, Appl
29	461	57.8	191	14	US-10-153-207-1	Sequence 1, Appli
30	461	57.8	191	14	US-10-400-377-1	Sequence 1, Appli
31	461	57.8	191	14	US-10-400-708-1	Sequence 1, Appli
32	461	57.8	191	14	US-10-298-148-1	Sequence 1, Appli
33	461	57.8	191	15	US-10-646-798-2	Sequence 2, Appli
34	461	57.8	191	15	US-10-621-693-2	Sequence 2, Appli
35	461	57.8	191	15	US-10-621-693-21	Sequence 21, Appl
36	461	57.8	191	15	US-10-621-693-80	Sequence 80, Appl
37	461	57.8	191	15	US-10-621-693-82	Sequence 82, Appl
38	461	57.8	191	15	US-10-621-693-84	Sequence 84, Appl
39	461	57.8	191	16	US-10-718-340-1	Sequence 1, Appli
40	461	57.8	191	16	US-10-658-834A-868	Sequence 868, App
41	461	57.8	191	16	US-10-658-834A-869	Sequence 869, App
42	461	57.8	191	16	US-10-658-834A-870	Sequence 870, App
43	461	57.8	191	16	US-10-658-834A-871	Sequence 871, App
44	461	57.8	191	16	US-10-658-834A-883	Sequence 883, App
45	461	57.8	191	16	US-10-658-834A-884	Sequence 884, App

#### ALIGNMENTS

##### RESULT 1

US-10-054-873-7

; Sequence 7, Application US/10054873

; Publication No. US20020164712A1

```

; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
;                      Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/CN98/00052
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

```

```

Query Match          100.0%; Score 797; DB 13; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.7e-78;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGSHLVEALYLVCGER 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGSHLVEALYLVCGER 120

Qy    121 GFFYTPKTRGIVEQCCTSICSLYQLENYCN 150
        ||||||||||||||||||||||||||||
Db    121 GFFYTPKTRGIVEQCCTSICSLYQLENYCN 150

```



RESULT 2  
 US-10-054-873-6  
 ; Sequence 6, Application US/10054873  
 ; Publication No. US20020164712A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gan, Zhong Ru  
 ; TITLE OF INVENTION: Chimeric Protein Containing an  
 ; Intramolecular Chaperone-Like Sequence  
 ;  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ;  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/054,873  
 ; FILING DATE: 22-Jan-2002  
 ; CLASSIFICATION: <Unknown>  
 ;  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO PCT/CN98/00052  
 ; FILING DATE: 31-MAR-1998  
 ; APPLICATION NUMBER: US 09/423,100  
 ; FILING DATE: 11-DEC-2000  
 ;  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mycroft, Frank J  
 ; REGISTRATION NUMBER: 46,946  
 ; REFERENCE/DOCKET NUMBER: 020167-000130US  
 ;  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 107 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: <Unknown>  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-10-054-873-6

Query Match 69.7%; Score 555.5; DB 13; Length 107;  
 Best Local Similarity 71.3%; Pred. No. 2.1e-52;  
 Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

Qy	1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP	60
Db	1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNP-----	49
Qy	61 TPSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLGSHLVEALYLVCGER	120
Db	50 -----LGTGPRFVNQHLGSHLVEALYLVCGER	77
Qy	121 GFFYTPKTRGIVEQCCTSICSLYQLENYCN	150

Db 78 GFFYTPKTRGIVEQCCTSICSLYQLENYCN 107

RESULT 3

US-10-054-873-2

; Sequence 2, Application US/10054873

; Publication No. US20020164712A1

; GENERAL INFORMATION:

; APPLICANT: Gan, Zhong Ru

; TITLE OF INVENTION: Chimeric Protein Containing an  
; Intramolecular Chaperone-Like Sequence

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/054,873

; FILING DATE: 22-Jan-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/CN98/00052

; FILING DATE: 31-MAR-1998

; APPLICATION NUMBER: US 09/423,100

; FILING DATE: 11-DEC-2000

; ATTORNEY/AGENT INFORMATION:

; NAME: Mycroft, Frank J

; REGISTRATION NUMBER: 46,946

; REFERENCE/DOCKET NUMBER: 020167-000130US

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 92 amino acids

; TYPE: amino acid

; STRANDEDNESS: <Unknown>

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-054-873-2

Query Match 59.0%; Score 470; DB 13; Length 92;

Best Local Similarity 100.0%; Pred. No. 3.4e-43;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

|||||

Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

Qy 61 TPSNREETQQKS NLELLRLISLLLIQSWLEPVQ 92

Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 4

US-09-819-094-24

; Sequence 24, Application US/09819094  
; Publication No. US20030186382A1  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Richard I.  
; APPLICANT: Martial, Joseph A.  
; APPLICANT: Struman, Ingrid  
; APPLICANT: Taylor, Robert  
; APPLICANT: Bentzien, Frauke  
; TITLE OF INVENTION: No. US20030186382A1el Antiangiogenic Peptide Agents and  
Their  
; TITLE OF INVENTION: Therapeutic and Diagnostic Use  
; FILE REFERENCE: UCSF-018/02US  
; CURRENT APPLICATION NUMBER: US/09/819,094  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 09/076,675  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: 60/046,394  
; PRIOR FILING DATE: 1997-05-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 24  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-819-094-24

Query Match 59.0%; Score 470; DB 10; Length 134;  
Best Local Similarity 100.0%; Pred. No. 5.6e-43;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
|  
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60  
  
Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92  
|  
Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 5

US-10-714-067-24

; Sequence 24, Application US/10714067  
; Publication No. US20040077054A1  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Richard I.  
; APPLICANT: Martial, Joseph A.  
; APPLICANT: Struman, Ingrid  
; APPLICANT: Taylor, Robert  
; APPLICANT: Bentzien, Frauke  
; TITLE OF INVENTION: Novel Antiangiogenic Peptide Agents and Their  
; TITLE OF INVENTION: Therapeutic and Diagnostic Use  
; FILE REFERENCE: UCSF-018/02US

```
; CURRENT APPLICATION NUMBER: US/10/714,067
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US/09/819,094
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/076,675
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/046,394
; PRIOR FILING DATE: 1997-05-12
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO 24
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-714-067-24
```

## RESULT 6

Query Match 58.5%; Score 466; DB 15; Length 188;  
Best Local Similarity 70.5%; Pred. No. 2.4e-42;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy      61 TPSNREETQQKSNLELLLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
        ||||||||||||||||||||||||||||| | | : | : |
Db      61 TPSNREETQQKSNLELLLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
        | : : | | | :|:| | :|
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

RESULT 7

US-09-819-094-23

; Sequence 23, Application US/09819094

; Publication No. US20030186382A1

; GENERAL INFORMATION:

; APPLICANT: Weiner, Richard I.

; APPLICANT: Martial, Joseph A.

; APPLICANT: Struman, Ingrid

; APPLICANT: Taylor, Robert

; APPLICANT: Bentzien, Frauke

; TITLE OF INVENTION: No. US20030186382A1el Antiangiogenic Peptide Agents and Their

; TITLE OF INVENTION: Therapeutic and Diagnostic Use

; FILE REFERENCE: UCSF-018/02US

; CURRENT APPLICATION NUMBER: US/09/819,094

; CURRENT FILING DATE: 2001-03-27

; PRIOR APPLICATION NUMBER: 09/076,675

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: 60/046,394

; PRIOR FILING DATE: 1997-05-12

; NUMBER OF SEQ ID NOS: 34

; SEQ ID NO 23

; LENGTH: 192

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-819-094-23

```

Query Match          58.5%; Score 466; DB 10; Length 192;
Best Local Similarity 70.5%; Pred. No. 2.4e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy      61 TPSNREETQQKSNLELLLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
        ||||||||||||||||||||||||||||| | | : | : |
Db      61 TPSNREETQQKSNLELLLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
        | : : | | | :|:| | :|
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

# RESULT 8

US-10-621-693-8

; Sequence 8, Application US/10621693

; Publication No. US20040059093A1

; GENERAL INFORMATION:

; APPLICANT: Gentide Biopharmaceuticals, Inc.

; APPLICANT: Bussell, Stuart

; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES AS

; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS

; FILE REFERENCE: GNT-00101.P.1-US

; CURRENT APPLICATION NUMBER: US/10/621,693

; CURRENT FILING DATE: 2003-07-16

; PRIOR APPLICATION NUMBER: US 60/396,466

; PRIOR FILING DATE: 2002-07-16

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8

; LENGTH: 192

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: synthetic sequence

; FEATURE:

; NAME/KEY: mat\_peptide

; LOCATION: (1)..()

US-10-621-693-8

Query Match 58.5%; Score 466; DB 15; Length 192;

Best Local Similarity 70.5%; Pred. No. 2.4e-42;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          |||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60

Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
          |||
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | | :|:| | :|
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

# RESULT 9

US-10-621-693-78

; Sequence 78, Application US/10621693

; Publication No. US20040059093A1

; GENERAL INFORMATION:

; APPLICANT: Gentide Biopharmaceuticals, Inc.

; APPLICANT: Bussell, Stuart

; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES AS

; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS

; FILE REFERENCE: GNT-00101.P.1-US

; CURRENT APPLICATION NUMBER: US/10/621,693

```
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 78
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
US-10-621-693-78
```

```
Query Match          58.5%; Score 466; DB 15; Length 192;
Best Local Similarity 70.5%; Pred. No. 2.4e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
```

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQLGTGPRFVNQHLCS-----HLV 110
          ||||||||||||||||||||||||||||| : | : |
Db      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | : | : |
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142
```

# RESULT 10

US-10-621-693-86

```
; Sequence 86, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Gentide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN
SEQUENCES AS
; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 86
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(192)
```





```

          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60
Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
          ||||||||||||||||||||||||||||| | | : | : |
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119
Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | | : | : | | : |
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

RESULT 12

```

US-10-621-693-42
; Sequence 42, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Gentide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN
SEQUENCES AS
; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
US-10-621-693-42

```

```

Query Match          58.5%; Score 466; DB 15; Length 193;
Best Local Similarity 70.5%; Pred. No. 2.4e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```

```

Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60
Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
          ||||||||||||||||||||||||||||| | | : | : | |
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119
Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | | : | : | | : |
Db      120 EGIQTLMGRLEDG---SPRTGQIFKQ 142

```

RESULT 13

```

US-10-621-693-72
; Sequence 72, Application US/10621693

```

```
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Gentide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN
SEQUENCES AS
; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
US-10-621-693-72
```

```
Query Match          58.5%; Score 466; DB 15; Length 206;
Best Local Similarity 70.5%; Pred. No. 2.7e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
```

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60
          |||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

Qy      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQLGTGPRFVNQHLCS-----HLV 110
          |||
Db      61 TPSNREETQQKSNELELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | | :|:| | :|
Db      120 EGIQTLMGRLDG---SPRTGQIFKQ 142
```

#### RESULT 14

US-10-621-693-51

```
; Sequence 51, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Gentide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN
SEQUENCES AS
; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
```

```
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1)..()
US-10-621-693-51
```

```
Query Match          58.5%; Score 466; DB 15; Length 391;
Best Local Similarity 70.5%; Pred. No. 6.1e-42;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;
```

```
Qy      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60
          |||||
Db      1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCS-----HLV 110
          |||||
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSILVYGASDSNVYDLLKDLE 119

Qy      111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | :!| | :|
Db      120 EGIQTIMGRLEDG---SPRTGQIFKQ 142
```

# RESULT 15

US-10-621-693-32

```
; Sequence 32, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Gentide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN
SEQUENCES AS
; TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; CURRENT FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (379)..(569)
; OTHER INFORMATION: sequence is repeated N-1 times, where N is a positive
whole numbe
; FEATURE:
; NAME/KEY: mat_peptide
```

; LOCATION: (1)..()  
US-10-621-693-32

Query Match 58.5%; Score 466; DB 15; Length 574;  
Best Local Similarity 70.5%; Pred. No. 1e-41;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

```
Qy      1 MFPTIPLSRLEFDNAMLRAHRLHQLAFD TYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60
          |||
Db      1 MFPTIPLSRLEFDNAMLRAHRLHQLAFD TYQEFEEAYIPKEQKYSFLQNPQTSLSFSSESIP 60

Qy     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCGS-----HLV 110
          |||
Db     61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLE 119

Qy    111 EALYLVCG--ERGFFYTPKTRGIVEQ 134
          | : : | | | :|:| | :|
Db    120 EGIQTLMGRLEDG---SPRTGQIFKQ 142
```

Search completed: February 11, 2005, 19:03:54  
Job time : 110.041 secs

OM protein - protein search, using sw model

Run on: February 11, 2005, 17:42:04 ; Search time 130.904 Seconds  
(without alignments)  
586.780 Million cell updates/sec

Title: US-10-054-873-7  
Perfect score: 797  
Sequence: 1 MEFTIPLSRLFDNAMLRAHR.....IVEQCCTSICSLYQLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

		8					
Result		Query					
No.	Score	Match Length DB	ID			Description	
1	461	57.8 217 1	SOMA_HUMAN			P01241 homo sapien	
2	461	57.8 217 1	SOMA_PANTR			P58756 pan troglod	
3	460	57.7 217 1	SOMA_MACMU			P33093 macaca mula	
4	460	57.7 217 2	Q6IYF0			Q6iyf0 homo sapien	
5	457	57.3 217 2	Q6IYF1			Q6iyf1 homo sapien	
6	437	54.8 217 1	SOMA_SAIIBB			P58343 saimiri bol	
7	436	54.7 217 2	Q8WNE0			Q8wne0 ateles geof	
8	434.5	54.5 217 1	SOM2_PANTR			P58757 pan troglod	
9	432	54.2 217 1	SOMA_CALJA			Q9gmb3 callithrix	
10	427.5	53.6 217 2	Q6FH54			Q6fh54 homo sapien	
11	427.5	53.6 245 2	O14644			O14644 homo sapien	
12	426.5	53.5 217 1	SOM2_HUMAN			P01242 homo sapien	
13	421.5	52.9 217 2	Q6FH32			Q6fh32 homo sapien	
14	407.5	51.1 217 2	Q07369			Q07369 macaca mula	
15	399	50.1 184 2	Q866T9			Q866t9 pan troglod	

16	399	50.1	217	1	SOM2_MACMU	Q07370	macaca mula
17	397	49.8	217	2	Q866U1	Q866u1	pan troglod
18	396	49.7	212	2	Q07368	Q07368	macaca mula
19	396	49.7	217	2	Q07367	Q07367	macaca mula
20	385	48.3	217	2	Q866T8	Q866t8	pan troglod
21	381	47.8	217	1	CSH_HUMAN	P01243	homo sapien
22	381	47.8	217	2	Q6PF11	Q6pf11	homo sapien
23	370	46.4	217	2	Q866U0	Q866u0	pan troglod
24	348	43.7	217	2	Q8WND9	Q8wnd9	ateles geof
25	341	42.8	202	2	O14643	O14643	homo sapien
26	322.5	40.5	217	2	Q8MI74	Q8mi74	callithrix
27	310.5	39.0	216	1	SOMA_MESAU	P37886	mesocricetu
28	307.5	38.6	190	1	SOMA_BALBO	P33092	balaenopter
29	307	38.5	184	2	Q69B30	Q69b30	ateles geof
30	306.5	38.5	216	1	SOMA_HORSE	P01245	equus cabal
31	306.5	38.5	216	2	O70615	O70615	spalax leuc
32	306.5	38.5	217	1	SOMA_GALSE	Q9gka1	galago sene
33	306.5	38.5	217	1	SOMA_NYCPY	Q9gmb2	nycticebus
34	304.5	38.2	216	1	SOMA_MOUSE	P06880	mus musculu
35	302.5	38.0	216	1	SOMA_RABIT	P46407	oryctolagus
36	302.5	38.0	216	1	SOMA_RAT	P01244	rattus norv
37	301.5	37.8	190	1	SOMA_LOXAF	P20392	loxedonta a
38	301.5	37.8	216	1	SOMA_BALPH	Q659q8	balaenopter
39	301.5	37.8	216	1	SOMA_CANFA	P33711	canis famil
40	301.5	37.8	216	1	SOMA_FELCA	P46404	felis silve
41	301.5	37.8	216	1	SOMA_PIG	P01248	sus scrofa
42	301.5	37.8	216	2	Q8HYE5	Q8hye5	ailuropoda
43	301.5	37.8	216	2	Q8MI73	Q8mi73	delphinus d
44	301.5	37.8	216	2	Q7YQB8	Q7yqb8	hippopotamu
45	299.5	37.6	216	1	SOMA_MUSVI	P19795	mustela vis

#### ALIGNMENTS

##### RESULT 1

##### SOMA\_HUMAN

ID SOMA\_HUMAN STANDARD; PRT; 217 AA.  
 AC P01241; Q14405; Q16631; Q9HBZ1; Q9UMJ7; Q9UNL5;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).  
 GN Name=GH1;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=80034477; PubMed=386281;  
 RA Roskam W., Rougeon F.;  
 RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."  
 RT Nucleic Acids Res. 7:305-320(1979).  
 RL Nucleic Acids Res. 7:305-320(1979).  
 RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=79203293; PubMed=377496;  
 RA Martial J.A., Hallewell R.A., Baxter J.D., Goodman H.M.;  
 RT "Human growth hormone: complementary DNA cloning and expression in  
 RT bacteria.";  
 RL Science 205:602-607(1979).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.  
 RX MEDLINE=82014939; PubMed=6269091;  
 RA Denoto F.M., Moore D.D., Goodman H.M.;  
 RT "Human growth hormone DNA sequence and mRNA structure: possible  
 RT alternative splicing.";  
 RL Nucleic Acids Res. 9:3719-3730(1981).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83182010; PubMed=7169009;  
 RA Seeburg P.H.;  
 RT "The human growth hormone gene family: nucleotide sequences show  
 RT recent divergence and predict a new polypeptide hormone.";  
 RL DNA 1:239-249(1982).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89307277; PubMed=2744760;  
 RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinas R.E.,  
 RA Seeburg P.H.;  
 RT "The human growth hormone locus: nucleotide sequence, biology, and  
 RT evolution.";  
 RL Genomics 4:479-497(1989).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Pituitary;  
 RA Gu J., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;  
 RT "A novel gene expressed in human pituitary.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM 4).  
 RC TISSUE=Pituitary;  
 RX MEDLINE=20402571; PubMed=10931946; DOI=10.1073/pnas.160270997;  
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
 RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
 RT axis and full-length cDNA cloning.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
 RN [8]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=86137393; PubMed=3912261; DOI=10.1016/0378-1119(85)90319-1;  
 RA Gray G.L., Baldridge J.S., McKeown K.S., Heyneker H.L., Chang C.N.;  
 RT "Periplasmic production of correctly processed human growth hormone in  
 RT Escherichia coli: natural and bacterial signal sequences are  
 RT interchangeable.";  
 RL Gene 39:247-254(1985).  
 RN [9]  
 RP SEQUENCE OF 27-217.  
 RX MEDLINE=69289202; PubMed=5810834;

RA Li C.H., Dixon J.S., Liu W.-K.;  
 RT "Human pituitary growth hormone. XIX. The primary structure of the  
 RT hormone.";  
 RL Arch. Biochem. Biophys. 133:70-91(1969).  
 RN [10]  
 RP SEQUENCE OF 27-217, AND REVISIONS.  
 RX MEDLINE=72143935; PubMed=5144027;  
 RA Li C.H., Dixon J.S.;  
 RT "Human pituitary growth hormone. 32. The primary structure of the  
 RT hormone: revision.";  
 RL Arch. Biochem. Biophys. 146:233-236(1971).  
 RN [11]  
 RP REVISION.  
 RX MEDLINE=73092028; PubMed=4675454;  
 RA Bewley T.A., Dixon J.S., Li C.H.;  
 RT "Sequence comparison of human pituitary growth hormone, human  
 RT chorionic somatomammotropin, and ovine pituitary growth and lactogenic  
 RT hormones.";  
 RL Int. J. Pept. Protein Res. 4:281-287(1972).  
 RN [12]  
 RP SEQUENCE OF 27-61 AND 102-124.  
 RX MEDLINE=71139765; PubMed=5279046;  
 RA Niall H.D.;  
 RT "Revised primary structure for human growth hormone.";  
 RL Nature New Biol. 230:90-91(1971).  
 RN [13]  
 RP REVISIONS TO 119-120 AND 157-159.  
 RX MEDLINE=71153968; PubMed=5279528;  
 RA Niall H.D., Hogan M.L., Sauer R., Rosenblum I.Y., Greenwood F.C.;  
 RT "Sequences of pituitary and placental lactogenic and growth hormones:  
 RT evolution from a primordial peptide by gene reduplication.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).  
 RN [14]  
 RP REVISION.  
 RA Niall H.D.;  
 RT "The chemistry of the human lactogenic hormones.";  
 RL (In) Griffiths K. (eds.);  
 RL Prolactin and carcinogenesis, Proc. fourth tenovus workshop prolactin,  
 RL pp.13-20, Alpha Omega Alpha Press, Cardiff (1972).  
 RN [15]  
 RP SEQUENCE OF 27-79 (ISOFORM 2).  
 RX MEDLINE=81117361; PubMed=7462247;  
 RA Chapman G.E., Rogers K.M., Brittain T., Bradshaw R.A., Bates O.J.,  
 RA Turner C., Cary P.D., Crane-Robinson C.;  
 RT "The 20,000 molecular weight variant of human growth hormone.  
 RT Preparation and some physical and chemical properties.";  
 RL J. Biol. Chem. 256:2395-2401(1981).  
 RN [16]  
 RP SEQUENCE OF 46-80 (ISOFORM 2).  
 RX MEDLINE=80130196; PubMed=7356479;  
 RA Lewis U.J., Bonewald L.F., Lewis L.J.;  
 RT "The 20,000-dalton variant of human growth hormone: location of the  
 RT amino acid deletions.";  
 RL Biochem. Biophys. Res. Commun. 92:511-516(1980).  
 RN [17]  
 RP DEAMIDATION OF GLN-163 AND ASN-178.  
 RX MEDLINE=82052997; PubMed=7028740;



RA Lewis U.J., Singh R.N., Bonewald L.F., Seavey B.K.;  
 RT "Altered proteolytic cleavage of human growth hormone as a result of  
 RT deamidation.";  
 RL J. Biol. Chem. 256:11645-11650(1981).  
 RN [18]  
 RP PHOSPHORYLATION SITES SER-132 AND SER-176.  
 RC TISSUE=Pituitary;  
 RX PubMed=14997482; DOI=10.1002/pmic.200300584;  
 RA Giorgianni F., Beranova-Giorgianni S., Desiderio D.M.;  
 RT "Identification and characterization of phosphorylated proteins in the  
 RT human pituitary.";  
 RL Proteomics 4:587-598(2004).  
 RN [19]  
 RP REVIEW.  
 RX MEDLINE=99321812; PubMed=10393484; DOI=10.1159/000053128;  
 RA Baumann G.;  
 RT "Growth hormone heterogeneity in human pituitary and plasma.";  
 RL Horm. Res. 51 Suppl. 1:2-6(1999).  
 RN [20]  
 RP 3D-STRUCTURE MODELING.  
 RX MEDLINE=88190073; PubMed=3447173;  
 RA Cohen F.E., Kuntz I.D.;  
 RT "Prediction of the three-dimensional structure of human growth  
 RT hormone.";  
 RL Proteins 2:162-166(1987).  
 RN [21]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=92196577; PubMed=1549776;  
 RA de Vos A.M., Ultsch M., Kossiakoff A.A.;  
 RT "Human growth hormone and extracellular domain of its receptor:  
 RT crystal structure of the complex.";  
 RL Science 255:306-312(1992).  
 RN [22]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=95075462; PubMed=7984244; DOI=10.1038/372478a0;  
 RA Somers W., Ultsch M., de Vos A.M., Kossiakoff A.A.;  
 RT "The X-ray structure of a growth hormone-prolactin receptor complex.";  
 RL Nature 372:478-481(1994).  
 RN [23]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RA Chantalat L., Chirgadze N.Y., Jones N., Korber F., Navaza J.,  
 RA Pavlovsk A.G., Wlodawer A.;  
 RT "The crystal-structure of wild-type growth-hormone at 2.5-A  
 RT resolution.";  
 RL Protein Pept. Lett. 2:333-340(1995).  
 RN [24]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=97113023; PubMed=8943276; DOI=10.1074/jbc.271.50.32197;  
 RA Sundstroem M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,  
 RA Norstedt G.;  
 RT "Crystal structure of an antagonist mutant of human growth hormone,  
 RT G120R, in complex with its receptor at 2.9-A resolution.";  
 RL J. Biol. Chem. 271:32197-32203(1996).  
 RN [25]  
 RP VARIANT KOWARSKI SYNDROME CYS-103.  
 RX MEDLINE=96150232; PubMed=8552145; DOI=10.1056/NEJM199602153340704;  
 RA Takahashi Y., Kaji H., Okimura Y., Goji K., Abe H., Chihara K.;

RT "Short stature caused by a mutant growth hormone.";  
 RL N. Engl. J. Med. 334:432-436(1996).  
 RN [26]  
 RP ERRATUM.  
 RA Takahashi Y., Kaji H., Okimura Y., Goji K., Abe H., Chihara K.;  
 RL N. Engl. J. Med. 334:1207-1207(1996).  
 RN [27]  
 RP VARIANT KOWARSKI SYNDROME GLY-138.  
 RX MEDLINE=97426478; PubMed=9276733;

Query Match 57.8%; Score 461; DB 1; Length 217;  
 Best Local Similarity 70.3%; Pred. No. 1.6e-36;  
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 86  
 QY 62 PSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPRFVNQHLCS-----HLVE 111  
 |||||||||||||||||||||||||||| | | : | : | |  
 Db 87 PSNREETQQKSNLELLRLISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDLLE 145  
 QY 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134  
 : : | | : | : | : |  
 Db 146 GIQTLMGRLDQ---SPRTGQIFKQ 167

# RESULT 2

## SOMA\_PANTR

ID SOMA\_PANTR STANDARD; PRT; 217 AA.  
 AC P58756;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).  
 GN Name=GH1;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;  
 RT "Independent duplication of the growth hormone gene in three  
 RT Anthropoidean lineages.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Plays an important role in growth control. Its major  
 CC role in stimulating body growth is to stimulate the liver and  
 CC other tissues to secrete IGF-1. It stimulates both the  
 CC differentiation and proliferation of myoblasts. It also stimulates  
 CC amino acid uptake and protein synthesis in muscle and other  
 CC tissues (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: Belongs to the somatotropin/prolactin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration





Q6IYF0

Query Match 57.7%; Score 460; DB 2; Length 217;  
Best Local Similarity 98.9%; Pred. No. 2e-36;  
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLSFSES IPT  61
          |||
Db      27  FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSES IPT  86
          |||

Qy      62  PSNREETQOKSNLELLRISLLLIQSWLEPVQ  92
          |||
Db      87  PSNREETQOKSNLELLRISLLLIQSWLEPVQ  117
          |||

```

Q6IYF1

```
ID Q6IYF1 PRELIMINARY; PRT; 217 AA.
AC Q6IYF1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Growth hormone 1 variant 1.
GN Name=GH1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
```

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jorge A.A.L., Arnhold I.J.P., Mendonca B.B.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY613431; AAT11508.1; -.  
 DR HSSP; P01241; 1A22.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0005179; F:hormone activity; IEA.  
 DR InterPro; IPR009079; 4\_helix\_cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 SQ SEQUENCE 217 AA; 24875 MW; 12DB1B92F63934D8 CRC64;

Query Match 57.3%; Score 457; DB 2; Length 217;  
 Best Local Similarity 69.7%; Pred. No. 3.8e-36;  
 Matches 101; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSES IPT 61  
 Db 27 FPTIPLSRLFDNVMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSSES IPT 86  
 Qy 62 PSNREETQQKSNLELLRLISLLLIQSWLEPVQLGTGPREFVNQHLCGS-----HLVE 111  
 Db 87 PSNREETQQKSNLELLRLISLLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLLKDL EE 145  
 Qy 112 ALYLVCG--ERGFFYTPKTRGIVEQ 134  
 Db 146 GIQTLMGRL EDG---SPRTGQIFKQ 167

# RESULT 6

## SOMA\_SAIBB

ID SOMA\_SAIBB STANDARD; PRT; 217 AA.  
 AC P58343;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Somatotropin precursor (Growth hormone).  
 GN Name=GHI;  
 OS Saimiri boliviensis boliviensis (Bolivian squirrel monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.  
 OX NCBI\_TaxID=39432;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21265430; PubMed=11371582;  
 RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;  
 RT "Episodic evolution of growth hormone in primates and emergence of the  
 RT species specificity of human growth hormone receptor."  
 RL Mol. Biol. Evol. 18:945-953(2001).  
 CC -!- FUNCTION: Plays an important role in growth control. Its major  
 CC role in stimulating body growth is to stimulate the liver and  
 CC other tissues to secrete IGF-1. It stimulates both the  
 CC differentiation and proliferation of myoblasts. It also stimulates

```

CC      amino acid uptake and protein synthesis in muscle and other
CC      tissues (By similarity).
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- SIMILARITY: Belongs to the somatotropin/prolactin family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF339060; AAK62287.1; -.
DR      HSSP; P01241; 1A22.
DR      InterPro; IPR009079; 4_helix_cytokine.
DR      InterPro; IPR001400; Somatotropin.
DR      Pfam; PF00103; Hormone_1; 1.
DR      PRINTS; PR00836; SOMATOTROPIN.
DR      PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR      PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW      Hormone; Pituitary; Signal.
FT      SIGNAL          1      26      By similarity.
FT      CHAIN           27     217     Somatotropin.
FT      DISULFID        79     191     By similarity.
FT      DISULFID       208     215     By similarity.
SQ      SEQUENCE      217 AA;  24864 MW;  9515289992C529F7 CRC64;

```

## RESULT 7

```
ID      Q8WNEO      PRELIMINARY;      PRT;      217 AA.
AC      Q8WNEO;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Growth hormone.
GN      Name=GH-N;
OS      Ateles geoffroyi (Black-handed spider monkey).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ateles.
OX      NCBI_TaxID=9509;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;
```

```

RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF374234; AAL72286.1; -.
DR HSSP; P01241; 1A22.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR Pfam; PF00103; Hormone_1; 1.
DR PRINTS; PR00836; SOMATOTROPIN.
DR PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
SO SEQUENCE 217 AA; 24894 MW; 425829FF41EEAAE6 CRC64;

```

Query Match 54.7%; Score 436; DB 2; Length 217;  
Best Local Similarity 66.9%; Pred. No. 4.2e-34;  
Matches 97; Conservative 8; Mismatches 24; Indels 16; Gaps 4;

```
Qy      2 FPTIPLSRLEFDNAMLRAHRLHQLAFTDYQEFE EAYIPKEQKYSFLQN PQTSLSFSESIPT 61  
        ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db     27 FPTIPLSRLLDNAMLRAHRLHQLAFTDYQEFE EAYIPKEQKYSFLQN PQTSLCFSESIPT 86  
  
Qy     62 PSNREETQQKSNLELLRISLLLIQSWLEPVQLGTGPRFVNQHLCG-----SHLVE 111  
        |: :: | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db     87 PASKKETQQKSNLELLRISLLLIQSWFEPVQF-LRSVFANSLLYGVSDSDVYEY LKDL EE 145  
  
Qy    112 ALYL VCG--ER GFF YTPKTRGIVEQ 134  
        : : | | | : | | |  
Db   146 GIOTLMGRLEDG---SPOTGEIFRO 167
```

## RESULT 8

SOM2 PANTR

```

ID      SOM2_PANTR      STANDARD;      PRT;      217 AA.
AC      P58757;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Growth hormone variant precursor (GH-V) (Placenta-specific growth
DE      hormone) (Growth hormone 2).
GN      Name=GH2;
OS      Pan troglodytes (Chimpanzee).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX      NCBI_TaxID=9598;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;
RT      "Independent duplication of the growth hormone gene in three
RT      Anthropeoidean lineages.";
RL      Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC      -!- FUNCTION: Plays an important role in growth control. Its major
CC      role in stimulating body growth is to stimulate the liver and
CC      other tissues to secrete IGF-1. It stimulates both the
CC      differentiation and proliferation of myoblasts. It also stimulates
CC      amino acid uptake and protein synthesis in muscle and other
CC      tissues.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC      -!- SIMILARITY: Belongs to the somatotropin/prolactin family.

```



CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF374233; AAL72285.1; -.  
 DR HSSP; P01241; 1A22.  
 DR InterPro; IPR009079; 4\_helix\_cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KW Glycoprotein; Hormone; Placenta; Signal.  
 FT SIGNAL 1 26 By similarity.  
 FT CHAIN 27 217 Growth hormone variant.  
 FT DISULFID 79 191 By similarity.  
 FT DISULFID 208 215 By similarity.  
 SQ SEQUENCE 217 AA; 24990 MW; 1592A429075677DE CRC64;

Query Match 54.5%; Score 434.5; DB 1; Length 217;  
 Best Local Similarity 78.9%; Pred. No. 5.8e-34;  
 Matches 90; Conservative 4; Mismatches 9; Indels 11; Gaps 1;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
 |||||:|||||  
 Db 27 FPTIPLSRLFDNAMLRAHRLYQLAYDTYQEFEEAYILKEQKYSFLQNPQTSLSFSESIPT 86  
 Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQL-----GTGPRFVNQHL 104  
 ||||:||||| | :||  
 Db 87 PSNRVKTQQKSNLELLRISLLLIQSWLEPVQLLRVVFANSLVYGASDSNVYRHL 140

# RESULT 9

SOMA\_CALJA  
 ID SOMA\_CALJA STANDARD; PRT; 217 AA.  
 AC Q9GMB3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Somatotropin precursor (Growth hormone).  
 GN Name=GHL;  
 OS Callithrix jacchus (Common marmoset).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.  
 OX NCBI\_TaxID=9483;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wallis O.C., Wallis M.;  
 RT "Cloning and characterisation of a putative growth hormone encoding  
 RT gene from the marmoset (Callithrix jacchus).";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Plays an important role in growth control. Its major





```
DR      InterPro; IPR001400; Somatotropin.
DR      Pfam; PF00103; Hormone_1; 1.
DR      PRINTS; PR00836; SOMATOTROPIN.
DR      PROSITE; PS00266; SOMATOTROPIN_1; 1.
KW      Signal.
FT      SIGNAL          1          26          Potential.
SO      SEQUENCE      245 AA;  27101 MW;  14CC7F8CD75D91C8 CRC64;
```

```

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT  61
        |||||||||||||||| ||:||:||:||:||:||||| ||||||||||||| |||||
Db      27  FPTIPLSRLFDNAMLRRRLYQLAYDTYQEFEEAYILKEQKYSFLQNPQTSLCFSESIPT  86

Qy      62  PSNREETQQKSNLELLRISLLLIQSWLEPVQL-----GTGPRFVNQHLCGSHLV  110
        |||| :||||||||||||||||||||||| | | :|| |
Db      87  PSNRVKTQQKSNLELLRISLLLIQSWLEPVQLLRSVFANSVLVGASDSNVYRHL--KDLE  144

Qy     111  EALYLVCG 118
        | : : |
Db     145  EGIOTLIG 152

```

RA Igout A., Scippo M.L., Franken F., Hennen G.;  
 RT "Cloning and nucleotide sequence of placental hGH-V cDNA.";  
 RL Arch. Int. Physiol. Biochim. 96:63-67(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89307277; PubMed=2744760;  
 RA Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinas R.E.,  
 RA Seeburg P.H.;  
 RT "The human growth hormone locus: nucleotide sequence, biology, and  
 RT evolution.";  
 RL Genomics 4:479-497(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP REVIEW.  
 RX MEDLINE=99321812; PubMed=10393484; DOI=10.1159/000053128;  
 RA Baumann G.;  
 RT "Growth hormone heterogeneity in human pituitary and plasma.";  
 RL Horm. Res. 51 Suppl. 1:2-6(1999).  
 CC -!- FUNCTION: Plays an important role in growth control. Its major  
 CC role in stimulating body growth is to stimulate the liver and  
 CC other tissues to secrete IGF-1. It stimulates both the  
 CC differentiation and proliferation of myoblasts. It also stimulates  
 CC amino acid uptake and protein synthesis in muscle and other  
 CC tissues.  
 CC -!- SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer, disulfide-  
 CC linked or non-covalently associated, in homopolymeric and  
 CC heteropolymeric combinations. Can also form a complex either with  
 CC GHBP or with the alpha2-macroglobulin complex.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=GH-V1;  
 CC IsoId=P01242-1; Sequence=Displayed;  
 CC Name=2; Synonyms=GH-V2;

```

CC      IsoId=P01242-2; Sequence=VSP_006203;
CC      Note=No experimental confirmation available;
CC      -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC      -!- SIMILARITY: Belongs to the somatotropin/prolactin family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; K00470; AAA98619.1; -.
DR      EMBL; J03756; AAB59547.1; -.
DR      EMBL; J03756; AAB59548.1; -.
DR      EMBL; M38451; AAA35891.1; -.
DR      EMBL; J03071; AAA52552.1; -.
DR      EMBL; BC020760; AAH20760.1; -.
DR      PIR; A28072; STHUV2.
DR      PIR; D32435; STHUV.
DR      HSSP; P01241; 1A22.
DR      Genew; HGNC:4262; GH2.
DR      H-InvDB; HIX0014077; -.
DR      MIM; 139240; -.
DR      GO; GO:0005179; F:hormone activity; TAS.
DR      InterPro; IPR009079; 4_helix_cytokine.
DR      InterPro; IPR001400; Somatotropin.
DR      Pfam; PF00103; Hormone_1; 1.
DR      PRINTS; PR00836; SOMATOTROPIN.
DR      PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR      PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW      Alternative splicing; Glycoprotein; Hormone; Placenta; Polymorphism;
KW      Signal.
FT      SIGNAL          1      26
FT      CHAIN           27      217      Growth hormone variant.
FT      DISULFID        79      191      By similarity.
FT      DISULFID       208      215      By similarity.
FT      CARBOHYD        166      166      N-linked (GlcNAc. . .) (Potential).
FT      VARSPLIC        153      217      RLEDGSPRTGQIFNQSYSKFDTKSHNDDALLKNYGLLYCFR
FT                                     KDMDKVETFLRIVQCRSVEGSCGF -> VRVAPGIPNPGAP
FT                                     LASRDWGEKHCCPLFSSQALTQENSPYSSFPLVNPPGLSLQ
FT                                     PGGEGGKWMNERGREQCPSAWPLLLFLHFAEAGRWQPPDWA
FT                                     DLQSVLQQV (in isoform 2).
FT                                     /FTId=VSP_006203.
FT      VARIANT         90      90      R -> W (in dbSNP:5389).
FT                                     /FTId=VAR_014591.
FT      CONFLICT        109      109      I -> T (in Ref. 2).-
SQ      SEQUENCE       217 AA;  24999 MW;  7B9324698E822F96 CRC64;

Query Match          53.5%;  Score 426.5;  DB 1;  Length 217;
Best Local Similarity 78.1%;  Pred. No. 3.5e-33;
Matches 89;  Conservative 4;  Mismatches 10;  Indels 11;  Gaps 1;

Qy      2  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      27  FPTIPLSRLFDNAMLRRRLYQLAYDTYQEFEEAYILKEQKYSFLQNPQTSLCFSESIPT 86

```







RX PubMed=15246530; DOI=10.1016/j.gene.2004.03.034;  
 RA Revol De Mendoza A., Esquivel Escobedo D., Martinez Davila I.,  
 RA Saldana H.;  
 RT "Expansion and divergence of the GH locus between spider monkey and  
 RT chimpanzee.";  
 RL Gene 336:185-193(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Revol A., Esquivel D.E., Barrera H.S.;  
 RT "The GH-PL locus a hot-point between human and chimpanzee genomes.";  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY146627; AAN84507.1; -.  
 DR HSSP; P01241; 1AXI.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0005179; F:hormone activity; IEA.  
 DR InterPro; IPR009079; 4\_helix\_cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 FT NON\_TER 184 184  
 SQ SEQUENCE 184 AA; 21145 MW; 68D1FF4AE59178DD CRC64;

Query Match 50.1%; Score 399; DB 2; Length 184;  
 Best Local Similarity 84.6%; Pred. No. 1.3e-30;  
 Matches 77; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61  
 |||||:||||| |||| |||||:||||| : ||| ||:||||  
 Db 27 FPTIPLSRLFDHAMLQAHRHQLAIDTYQEFEEAYIPKDQKYSFLHDSQTSFCFSDSIPT 86  
 QY 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92  
 ||| |||||:|||||:  
 Db 87 PSNMEETQQKSNLELLRISLLLIQSWLEPVR 117

Search completed: February 11, 2005, 18:22:50  
 Job time : 131.904 secs